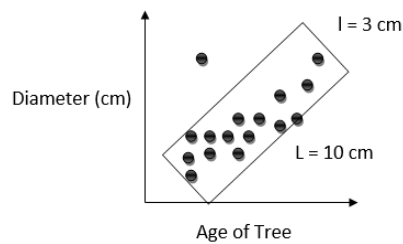


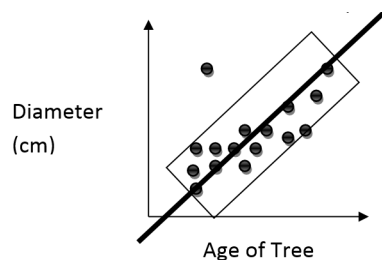
Regression Line or Line of Best-fit

It describes the **approximate** rule that the data should follow.

Ex.



After drawing a rectangle around the scatter plot (not including outliers), we will draw a straight line that best fits the direction of the points:



The line should:

- pass through the middle of the rectangle
- be parallel to its sides, and
- be as close to as many points as possible.

Nov 24-1:05 PM

Next, determine the equation of the line by choosing two points that the line passes through. (These points can be part of the raw data or read from the graph).

RECALL: to find the equation of a line passing through two points...

Step 1: Label the points - P1 (x_1, y_1) and P2 (x_2, y_2)

Step 2: Find the slope (a), using the formula:

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

Step 3: Find the y-intercept (b), using one of the points and the equation of a line:

$$y = ax + b$$

Nov 24-2:23 PM

What is the purpose of this line?

Predicting a value that isn't part of the data (extra-pollating), according to the line of best-fit by:

- Extending the line of best-fit and reading from the graph

OR

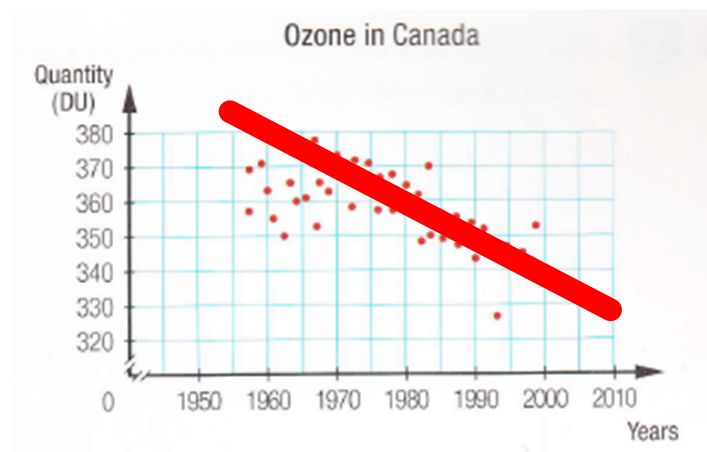
- Using the equation of the line and plugging in the value given for x or for y

Putting it all together:

A regression line allows you to predict the value(s) of one of the variables based on the value(s) of the other, and the correlation coefficient allows you to determine the reliability of this prediction.

Nov 24-2:49 PM

Ex. P. 115 #7



Nov 24-2:58 PM

Ex. P. 115 #7

Ozone in Canada

Quantity (DU)

Years

$y = ax + b$

$r \approx -\left(1 - \frac{\text{Short}}{\text{Long}}\right)$

$1 - \frac{2.6}{6.7}$

≈ -0.61

① $(x_1, y_1) = (1960, 370)$ $(x_2, y_2) = (1985, 340)$

② $a = \frac{340 - 370}{1985 - 1960} = \frac{-30}{25} = -1.2$

③ $y = ax + b$

$370 = -1.2(1960) + b$

$370 = -2352 + b$

+2352

$2722 = b$

$b = y_1 - ax_1$

$= 370 - (-1.2)(1960)$

$= 370 + 2352$

$= 2722$

④ $y = -1.2x + 2722$

For 2010:

$y = -1.2(2010) + 2722$

$= -2412 + 2722$

$= 310$

In 2016:

$y = -1.2(2016) + 2722$

$= 302.8 \text{ DU}$

Nov 24-2:58 PM

Homework

Textbook1:

P. 114 #5 & 6

P. 126 #4

Nov 24-3:39 PM