Close this window

Ask a Tutor

Review Index

Line of Best Fit(Least Square Method)

A **line of best fit** is a straight line that is the best approximation of the given set of data.

It is used to study the nature of the relation between two variables.

A line of best fit can be roughly determined using an eyeball method by drawing a straight line on a scatter plot so that the number of points above the line and below the line is about equal (and the line passes through as many points as possible).

A more accurate way of finding the line of best fit is the **least square method** .

Use the following steps to find the equation of line of best fit for a set of ordered pairs.

Step 1: Calculate the mean of the x-values and the mean of the y-values.

Step 2: Compute the sum of the squares of the x-values.

Step 3: Compute the sum of each x-value multiplied by its corresponding y-value.

Step 4: Calculate the slope of the line using the formula:

$$m = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

where n is the total number of data points.

Step 5: Compute the y-intercept of the line by using the formula:

$$b = \bar{y} - m\bar{x}$$

where $\bar{y}$ and $\bar{x}$ are the mean of the x- and y-coordinates of the data points respectively.

Step 6: Use the slope and the y -intercept to form the equation of the line.

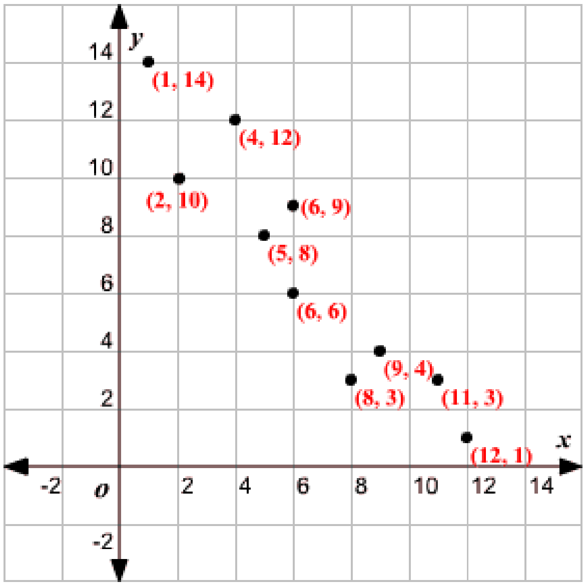
Example:

Use the least square method to determine the equation of line of best fit for the data. Then plot the line.

x	8	2	11	6	5	4	12	9	6	1
y	3	10	3	6	8	12	1	4	9	14

Solution:

Plot the points on a coordinate plane.



Calculate the means of the x-values and the y-values, the sum of squares of the x-values, and the sum of each x-value multiplied by its corresponding y-value.

x	y	xy	x^2
8	3	24	64
2	10	20	4
11	3	33	121
6	6	36	36
5	8	40	25
4	12	48	16
12	1	12	144
9	4	36	81
6	9	54	36
1	14	14	1
$\sum x = 64$	$\sum y = 70$	$\sum xy = 317$	$\sum x^2 = 528$

Calculate the slope.

$$m = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$= \frac{317 - \frac{(64)(70)}{10}}{528 - \frac{(64)^2}{10}}$$

$$\approx -1.1$$

Calculate the y-intercept.

First, calculate the mean of the x-values and that of the y-values.

$$\bar{x} = \frac{\sum x}{n} = \frac{64}{10} = 6.4 \qquad \bar{y} = \frac{\sum y}{n} = \frac{70}{10} = 7.0$$

Use the formula to compute the y-intercept.

$$\begin{aligned} b &= \bar{y} - m\bar{x} \\ &= 7.0 - (-1.1 \times 6.4) \\ &= 7.0 + 7.04 \\ &\approx 14.0 \end{aligned}$$

Use the slope and y-intercept to form the equation of the line of best fit.

The slope of the line is –1.1 and the y -intercept is 14.0.

Therefore, the equation is $y = -1.1\,x + 14.0$.

Draw the line on the scatter plot.