

Scatter Plots with Linear Associations



Objective

In this lesson, you will

informally fit lines to model data in scatter plots.

Creating Lines of Best Fit

A scatter plot consists of a group of data **points** that show the relationship between two quantities.

- If a line drawn through the middle of the data happens to be a straight line, the scatter plot shows a **linear** association.
- If a line drawn through a set of data is a **curved** line, the relationship is **nonlinear**.

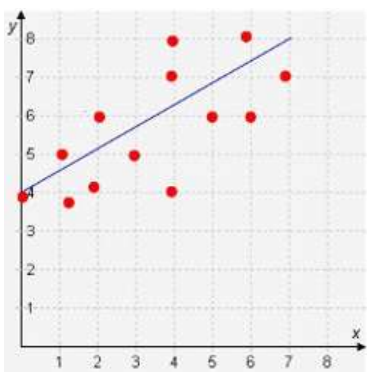
A line of best fit has data points **equally** distributed on either side of the line.

Every scatter plot can have a line of best fit, but some lines fit better than others.

So it's useful to describe how well the line of best fit describes the scatter plot.

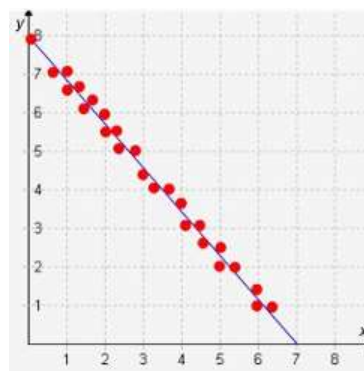
<p>moderate linear association</p>	<p>strong linear association</p>	<p>weak linear association</p>
<ul style="list-style-type: none"> → The line drawn through the data points is a straight line. → The points are loosely spread along the line. 	<ul style="list-style-type: none"> → The line drawn through the data points is a straight line. → The points are tightly clustered along the line. 	<ul style="list-style-type: none"> → The line drawn through the data points is a straight line. → The points are spread away from the line.

A relation's trend is also important.



positive association

- The line of best fit goes upward as it moves to the right.
- When one quantity increases, the other quantity increases.



negative association

- The line of best fit goes downward as it moves to the right.
- When one quantity increases, the other quantity decreases.

Equations of Lines of Best Fit

Once we've created the line of best fit, we can find the equation that models the relationship between the two quantities.



The slope-intercept form for the equation of a line is $y = mx + b$ where m is the slope of the line and b is the y-intercept.

$$\text{slope} = \frac{\text{rise}}{\text{run}}$$

Example:

In the graph:

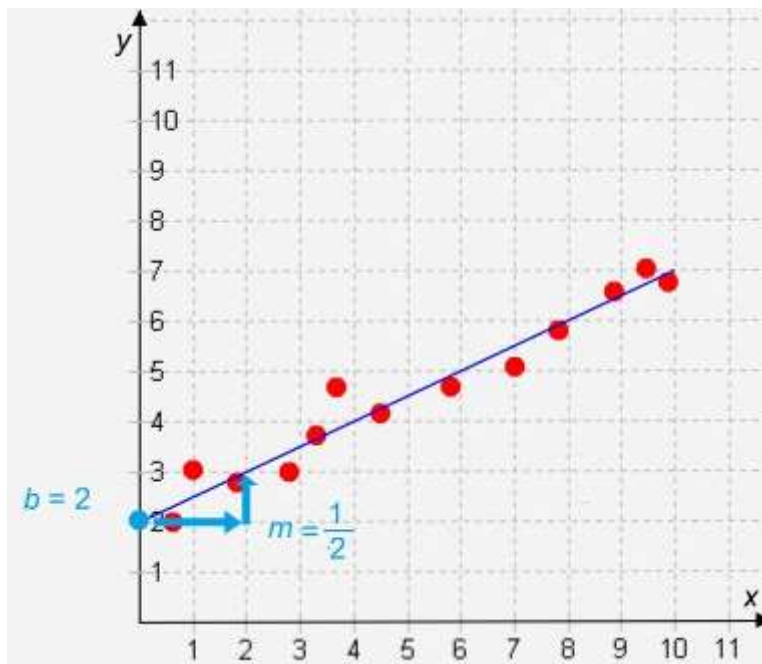
→ The slope is $\frac{1}{2}$.

→ The y-intercept is 2 .

(the y-coordinate when the x-coordinate is 0).

Substitute the values of m and b to get the

equation of the line in the graph: $y = \frac{1}{2}x + 2$.



Lesson Activity

Steven collects data on the height and weight of 20 dogs and plots the scatter plot shown.

Of lines *l*, *m*, and *n*, line l best fits the data.

If the line of best fit passes through the points (0, 0) and (30, 50), what is the equation of the line in the form $y = mx + b$?

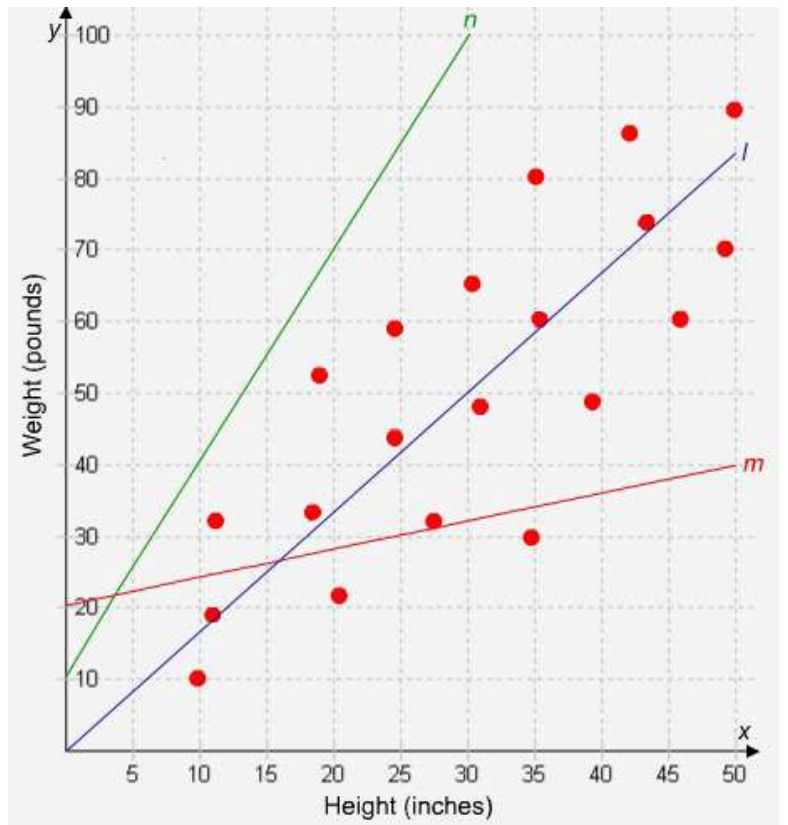
The value of y is 0 when the $x = 0$.

So, the y -intercept, b , is 0.

The slope, m , is $\frac{\boxed{50} - \boxed{0}}{\boxed{30} - \boxed{0}} = \frac{\boxed{50}}{\boxed{30}} = \frac{5}{\boxed{3}}$.

So, the equation of the line of best fit, passing through the points (0, 0) and (30, 50) is

$$y = \frac{5}{\boxed{3}}x + \boxed{0}$$



The equation of the line of best fit can be used to estimate the weight of a dog if its height is known.

Example:

Height in Inches (x)	Weight in Pounds (y)
12	20
15	25
18	30

← $\frac{5}{3}(12) + 0$

← $\frac{5}{3}(15) + 0$

← $\frac{5}{3}(18) + 0$

We can adjust the line to better fit the data.

A line of best fit has about the same number of data points **above** and **below** the line.

If this is not the case, we can change its **y**-intercept or its slope. Sometimes we may have to change both.

The graphs show the relationship between the age and the weight of puppies.

<p>This line seems to have the correct <input checked="" type="radio"/> slope <input type="radio"/> y-intercept, but the <input type="radio"/> slope <input checked="" type="radio"/> y-intercept is too high.</p>	<p>This line seems to have the correct <input type="radio"/> slope <input checked="" type="radio"/> y-intercept, but the <input checked="" type="radio"/> slope <input type="radio"/> y-intercept is not steep enough.</p>	<p>By adjusting both the y-intercept and the slope, we get this line of best fit.</p>

Summary

How does the equation of a line of best fit reveal whether there is a positive or negative association between the two quantities?

answers will vary