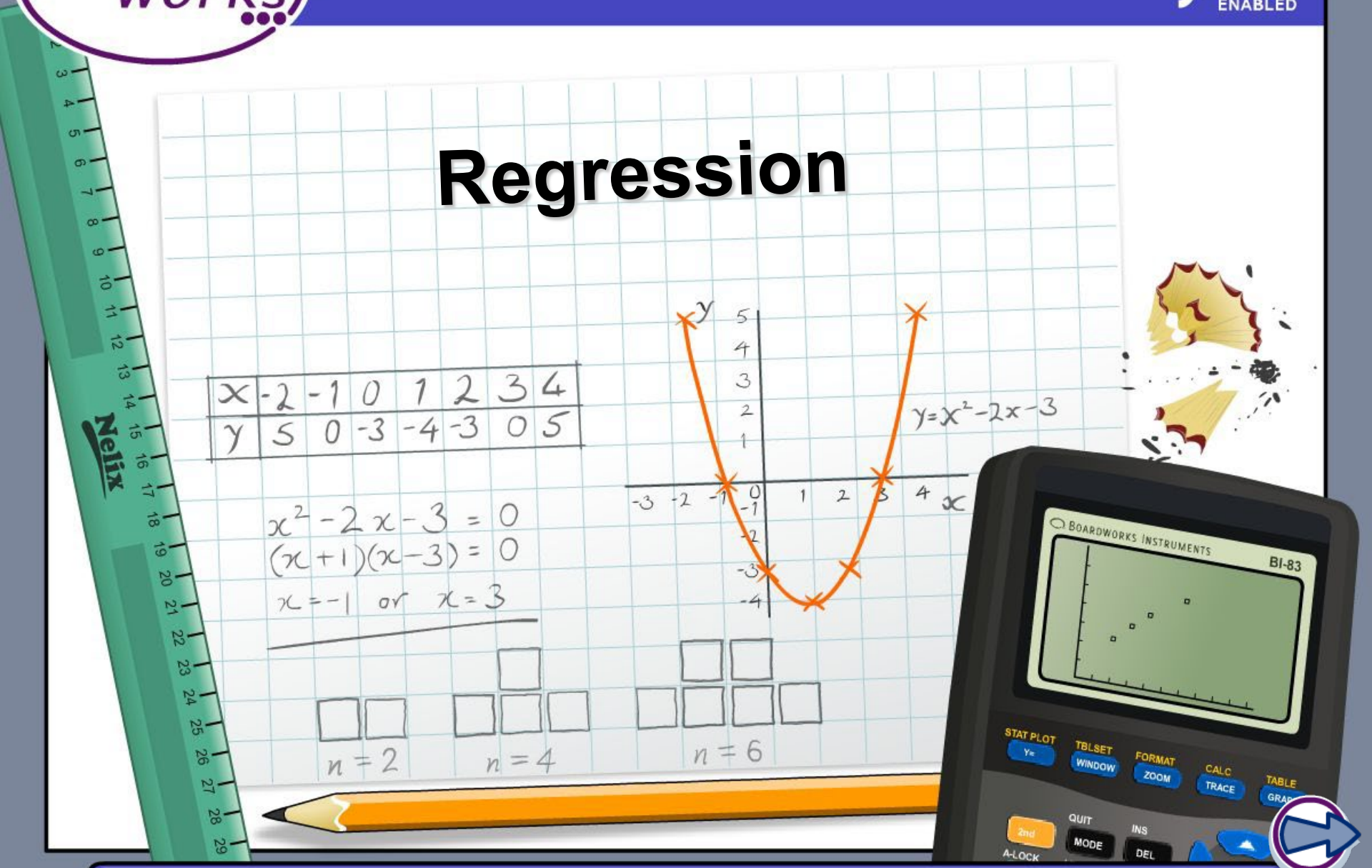
Regression



Information

Common core icons

This icon indicates a slide where the Standards for Mathematical Practice are being developed. Details of these are given in the Notes field.

Slides containing examples of mathematical modeling are marked with this stamp.

This icon indicates an opportunity for discussion or group work.

The **Standards for Mathematical Practice** outlined in the Common Core State Standards for Mathematics describe varieties of expertise that mathematics educators at all levels should seek to develop in their students.

They are:

1) Make sense of problems and persevere in solving them.

2) Reason abstractly and quantitatively.

3) Construct viable arguments and critique the reasoning of others.

4) Model with mathematics.

5) Use appropriate tools strategically.

6) Attend to precision.

7) Look for and make use of structure.

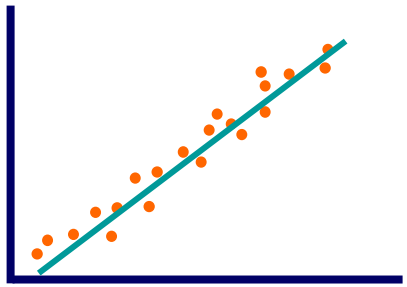
8) Look for and express regularity in repeated reasoning.

This icon indicates that the slide contains activities created in Flash. These activities are not editable.

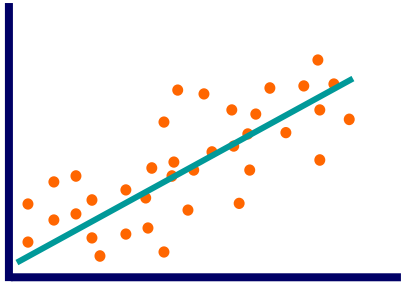
This icon indicates teacher's notes in the Notes field.

Lines of best fit

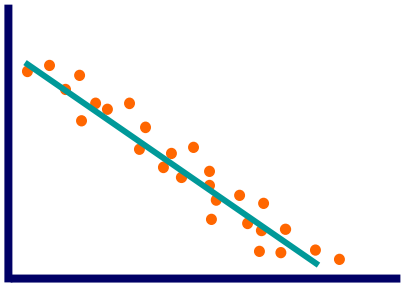
A **line of best fit** is drawn on a scatter plot to show the linear trend in a set of paired data.



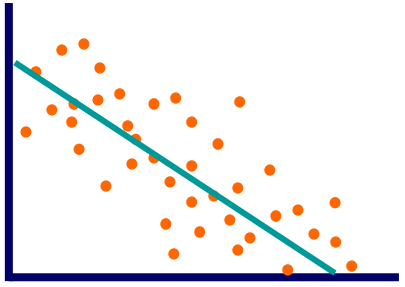
strong positive correlation



weak positive correlation



strong negative correlation



weak negative correlation

The stronger the correlation, the closer the points are to the line.

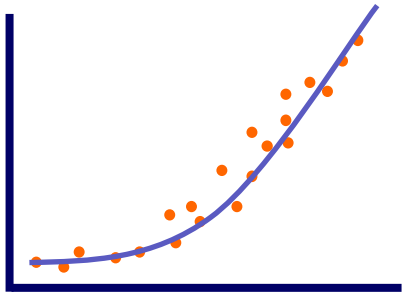
We can find the **equation** of a line of best fit by calculating the slope and y-intercept.

Choosing a function type

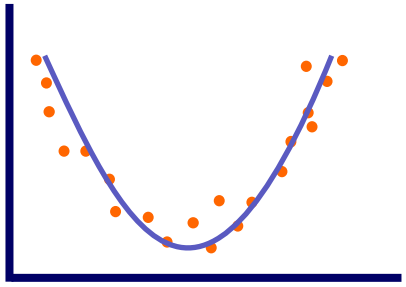
When we find the equation of the line of best fit for a scatter plot, we are effectively saying "This data follows a linear trend, so I can **fit a linear function** to the data."

Fitting a function to data in this way is called **regression** and the line we find is called the **regression line**.

Not all data follow a linear trend though.



exponential



quadratic

What type of function best fits these scatter plots?

What type of function?

What type of function would best fit each data set?
Plot the data from each table to identify the trend.
Press each table of values to see the answer.

a)

x	y
2	2
4	5
6	6
8	7
10	11
12	12

b)

x	y
0	30
1	25
2	18
3	13
4	17
5	24

c)

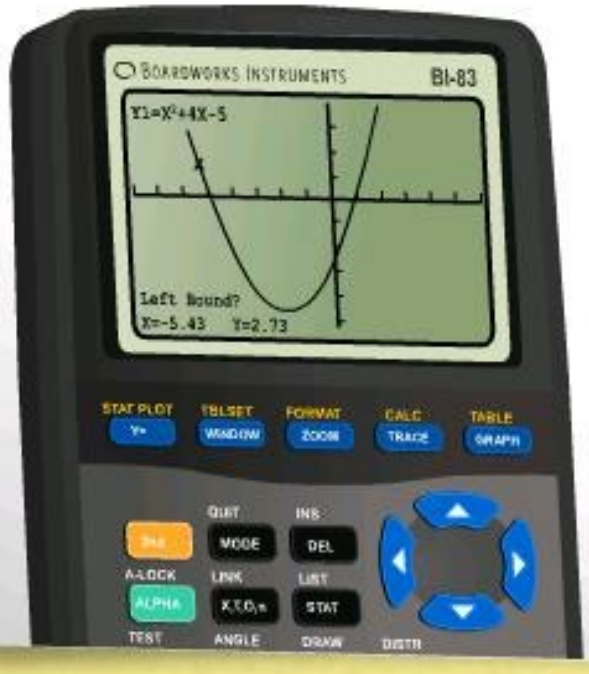
x	y
10	69
20	47
30	35
40	27
50	16
60	11

d)

x	y
0	1
5	6
10	10
15	19
20	28
25	39

Regression using calculators

linear | exponential | quadratic



You can use your graphing calculator to help you find the regression line for a set of paired data.

This is done using either the **LinReg**, **QuadReg** or **ExpReg** feature.

Press on each of the tabs above to find out about each type.

The value of r

When you use the linear regression feature, your calculator will display the value of r, the **correlation coefficient**.

The value of r indicates the strength of association between two variables. It shows how close points lie to the regression line.

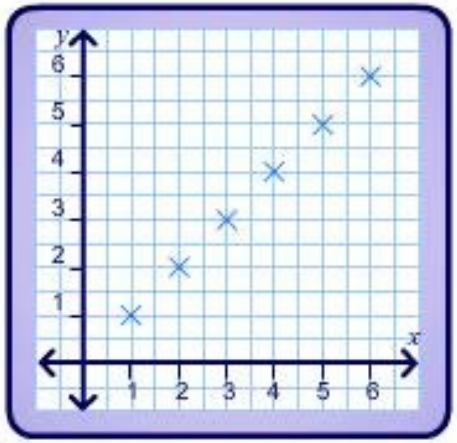
r can be between 1 and –1, inclusive

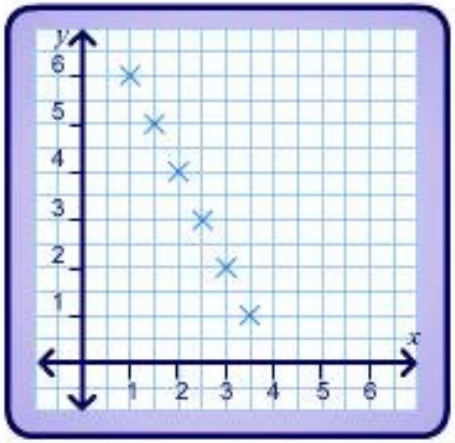
The closer the value of r is to 1 or –1, the better the regression line fits the data and the stronger the correlation.

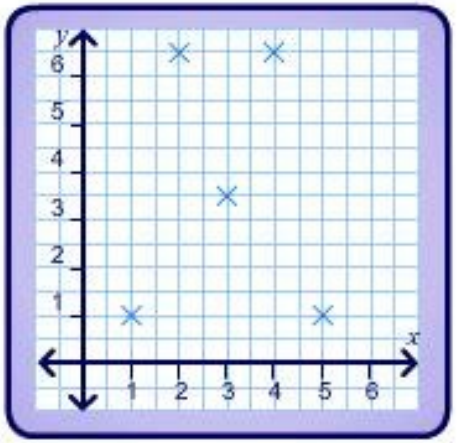
As r approaches zero from either side, it indicates that the equation is not a good fit for the data.

Linear correlation

Match each graph to the correct correlation descriptions.







$r = 0$

$r = -1$

$r = 1$

perfect linear negative correlation

no correlation

perfect linear positive correlation

App downloads

The table below shows the average number of apps downloaded per year from a popular app store since its launch in 2008.

Fit a function to the data and use it to predict the average number of downloads from the store in 2014.

year	2008	2009	2010	2011
average downloads (billions)	0.1	1.43	5.18	9.95

Press **play** to work through the solution.

Estimating values

The table below shows the average value of a particular model of car, depending on its age.

age (years)	2	4	6	8	10	12
value ($)	24,340	18,290	12,530	8,760	5,510	2,780

Use your graphing calculator to find a regression line for this data. Use it to estimate the original value of the car.

Plotting the data shows that it follows a linear pattern.

The regression equation is $y = -2141.5714x + 27026$.

The original value of the car will be at 0 years, which is the y-intercept of the regression line.

The value of the car was around **$27,026**.