

Height v. Shoe Size




Write down your height in inches.

Write down your shoe size.

X	Y
63	8.5
68	10.5
70	13.5
67	11
61	8
64	4.5
65	6.5
67	8
72	11.5
64	7
64	8
67	9.5
65	6.5
64	12
66	6.5
66	6.5
68	9
64	7.5
65	9
65	5.5
65	9.5
65	11.5

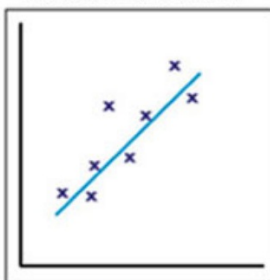
Scatter Plots on Desmos

- In row 1, Press “+” in the top right corner. Enter x values in x_1 and enter y values in y_1 
- Click the magnifying glass to zoom in toward the scatterplot.

Relationship

The relationship between a set of data is called **correlation**. A scatterplot can have positive, negative, or no correlation.

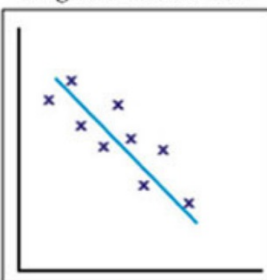
Positive correlation



The points lie close to a straight line, which has a positive gradient.

This shows that as one variable **increases** the other **increases**.

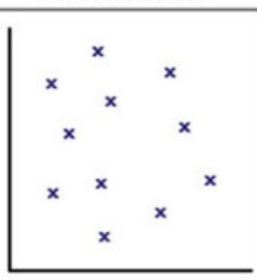
Negative correlation



The points lie close to a straight line, which has a negative gradient.

This shows that as one variable **increases**, the other **decreases**.

No correlation



There is no pattern to the points.

This shows that there is **no connection** between the two variables.

What correlation do we have?

A trend line for a scatterplot is called a **line of best fit** and it can be written in slope-intercept form.

Select two points from our data set.

What is the equation of the the line that passes through both points?

Line of Best Fit:

$$m = \underline{0.48} \quad b = \underline{-23.42}$$

$$y = 0.48x - 23.42$$

Meaning of slope: Your shoe size increases by 0.48 as your height increases by 1 inch.

Correlation Coefficient:

$$r = 0.51$$

Description: weak positive

Making Predictions

$$y = 0.48x - 23.42$$

What is the height (x) of a person who wears a size 15?

$$\begin{array}{r} 15 = 0.48x - 23.42 \\ + 23.42 \qquad + 23.42 \\ \hline 38.42 = 0.48x \\ 0.48 \qquad 0.48 \end{array}$$

$$x \approx 80 \text{ in.}$$

What is the shoe size (y) of a person who is 73 inches tall?

$$\begin{array}{l} y = 0.48(73) - 23.42 \\ y = 11.6 \end{array}$$

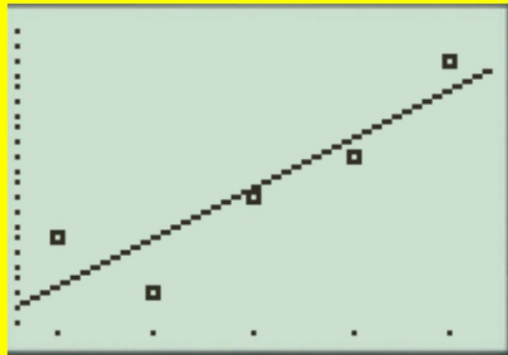
Enter the data into Desmos.

Hours Spent Studying (x)	Test Score
2	85
3	81
4	88
5	91
6	98

```
LinReg  
y=ax+b  
a=3.6  
b=74.2  
r2=.784503632  
r=.8857220963
```

Line of Best Fit

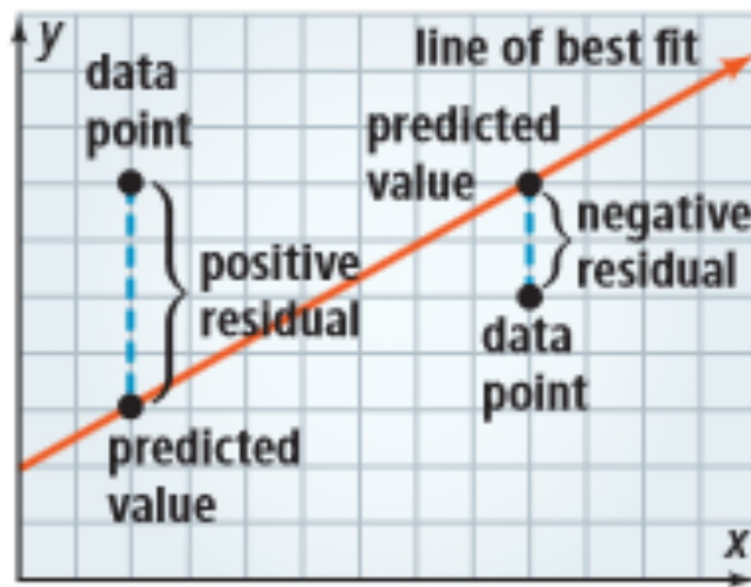
y = _____



Do the actual values measure perfectly to the predicted values (line of best fit)?

NO! The distance between the line of best fit and the actual values is called a **residual**.

$$\text{Residual} = \text{actual value} - \text{predicted value}$$



Residual - Vertical distance from actual value to predicted value.

Residuals have numerical values!

STAT--> EDIT--> L3

L1	L2	$\frac{\square}{\square}$	3
2	85		
3	81		
4	88		
5	91		
6	98		
-----	-----		
L3 =			

2nd--> STAT--> 7: RESID

NAME	OPS	MATH
1		L1
2		L2
3		L3
4		L4
5		L5
6		L6
7		RESID

L1	L2	$\frac{\square}{\square}$	3
2	85		
3	81		
4	88		
5	91		
6	98		
-----	-----		
L3 = LRESID			

Press ENTER

L1	L2	L3	3
2	85	8.7	
3	81	-4	
4	88	-6	
5	91	-1.2	
6	98	2.2	
-----	-----	-----	

The numbers in L3 represent the residuals for the (x_1, y_1) points in L1 and L2.

Positive Residuals v. Negative Residuals

Positive Residuals = the predicted is value **less than** the actual value.

For example, using the equation _____ when $x = 2$
 $y =$ _____.

This was _____ units **less than** the value from the table.

Negative Residuals = the predicted value is **greater than** the actual value.

Residual Plots

A residual plot is a graph that shows the residuals on the vertical axis and the independent variable on the horizontal axis. If the points in a residual plot are randomly dispersed around the horizontal axis, a linear regression model is appropriate for the data; otherwise, a non-linear model is more appropriate.

