

Algebra 1B7.7 Residuals

Residuals help to determine if a curve (shape) is appropriate for the data.  
(linear versus non-linear)

A residual is the difference between what is plotted in your scatter plot at a specific point, and what the regression equation predicts "should be plotted" at this specific point. If the scatter plot and the regression equation "agree" on a  $y$ -value (no difference), the residual will be zero.

**Definition:**

Residual = Observed  $y$ -value - Predicted  $y$ -value

A residual is the difference between the observed  $y$ -value (from scatter plot) and the predicted  $y$ -value (from regression equation line).

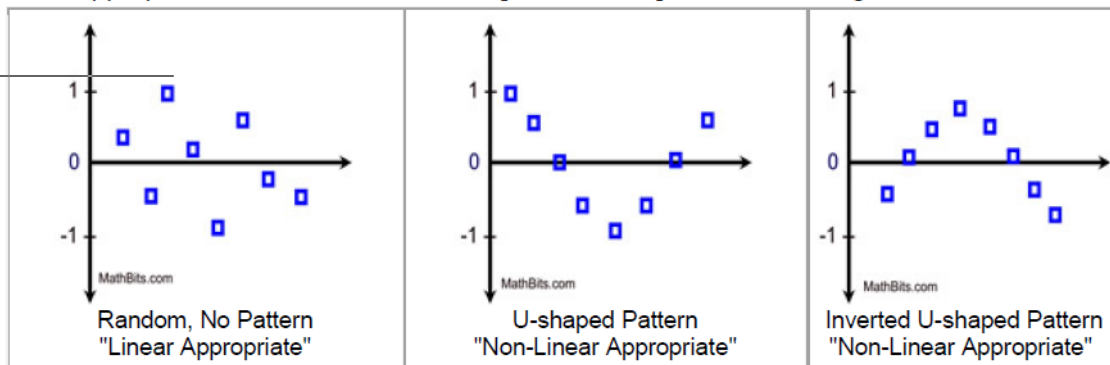
It is the vertical **distance** from the actual plotted point to the point on the regression line.

You can think of a residual as how far the data "fall" from the regression line  
(sometimes referred to as "*observed error*").

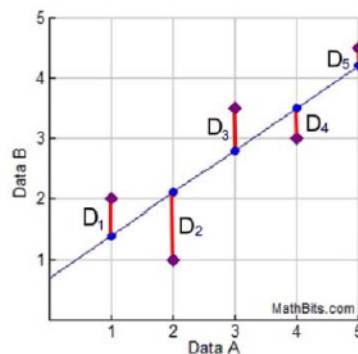
A **residual plot** is a scatter plot that shows the residuals on the vertical axis and the independent variable on the horizontal axis. The plot will help you to decide on whether a **linear model** is appropriate for your data.

Appropriate linear model: when plots are randomly placed, above and below  $x$ -axis ( $y = 0$ ).

Appropriate non-linear model: when plots follow a pattern, resembling a curve.



When a **pattern** is observed in a residual plot, a **linear regression model** is probably **not appropriate** for your data.



◆ Scatter Plot Points:

$\{(1,2), (2,1), (3,3\frac{1}{2}), (4,3), (5,4)\}$

● Regression Points

$\{(1,1.4), (2,2.1), (3,2.8), (4,3.5), (5,4.2)\}$

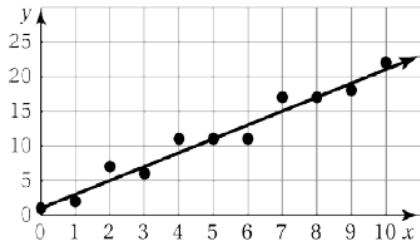
**The Red Line Segments:**

The red line segments represent the distances between the  $y$ -values of the actual scatter plot points, and the  $y$ -values of the regression equation at those points.

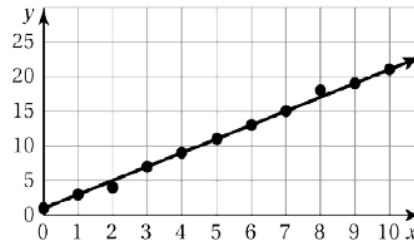
The lengths of the red line segments are called **RESIDUALS**.

**Example 1:** Compare and contrast the lines of fit for the two data sets below.

Graph 1:



Graph 2:



Describe what the line of best fit does for a set of data.

It is the best line that represents the data and allows the user to predict y-values.

What data set differs the most from the line of best fit?

The first one differs the most from the line that best fit because the data points are farther away from the line.

Calculate each graph's residual.

Graph 1:  $y = 2x + .5$

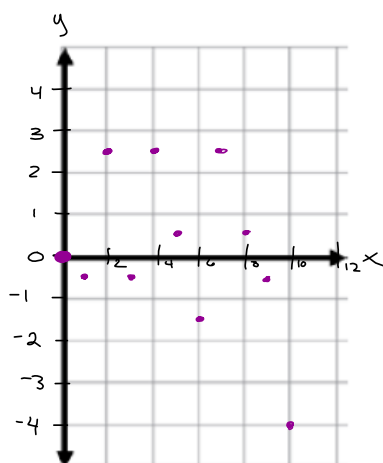
x	$y_{obs}$	$y_{pred}$	Residual $y_{obs} - y_{pred}$
0	1	1	$1 - 1 = 0$
1	2	2.5	$2 - 2.5 = -0.5$
2	7	4.5	$7 - 4.5 = 2.5$
3	6	6.5	$6 - 6.5 = -0.5$
4	11	8.5	$11 - 8.5 = 2.5$
5	11	10.5	$11 - 10.5 = 0.5$
6	11	12.5	-1.5
7	17	14.5	2.5
8	17	16.5	0.5
9	18	18.5	-0.5
10	22	26	-4

Graph 2:  $y = 2x + .5$

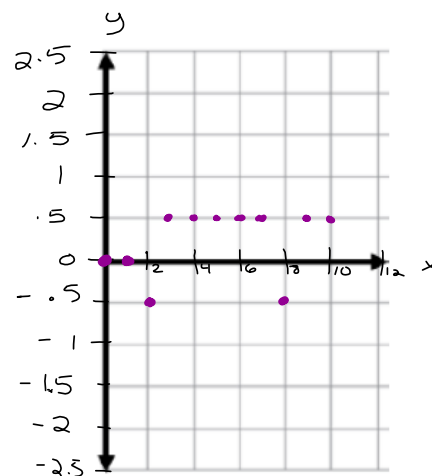
x	$y_{obs}$	$y_{pred}$	Residual $y_{obs} - y_{pred}$
0	1	1	0
1	2.5	2.5	0
2	4	4.5	-0.5
3	7	6.5	0.5
4	9	8.5	0.5
5	11	10.5	0.5
6	13	12.5	0.5
7	15	14.5	0.5
8	16	16.5	-0.5
9	19	18.5	0.5
10	21	<del>26</del> 20.5	0.5

Create Residual Plot (a scatterplot with those numbers the x-values and y residual values.) Describe what the two residual plots tell you about the data.

Graph 1:



Graph 2:



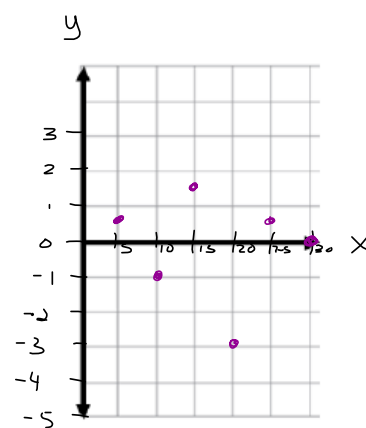
Practice:

Complete each table using the given linear regression (Round answers to one decimal place).  
Construct a residual plot.

1. Linear regression equation:  $y = 0.5x$ 

Pobserve

x	y	Predicted Value	Residual Value
5	3	2.5	0.5
10	4	5	-1
15	9	7.5	1.5
20	7	10	-3
25	13	12.5	0.5
30	15	15	0

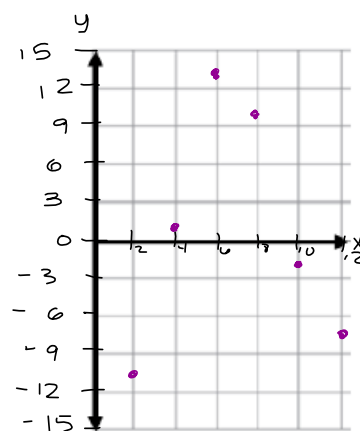


Does the residual plot suggest a linear relationship? Explain.

Yes, No Curved Pattern

2. Linear regression equation:  $y = -0.4x + 16.3$

x	y	Predicted Value	Residual Value
2	5	15.5	-10.5
4	15	14.7	0.3
6	26	13.9	12.1
8	23	13.1	9.9
10	11	12.3	-1.3
12	3	11.5	-8.5



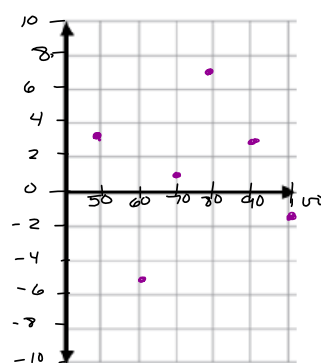
Does the residual plot suggest a linear relationship? Explain.

No, there is a curved (u-shaped) pattern

3. Linear regression equation:  $y = 4.9x + 16.4$

x	y	Predicted Value	Residual Value
100	505	506.4	-1.4
90	460	457.4	2.6
80	415	408.4	6.6
70	360	359.4	0.6
60	305	310.4	-5.4
50	265	261.4	3.6

be careful  
is order  
reversed



Does the residual plot suggest a linear relationship? Explain.

Yes, No curved pattern