Guided Practice 4.2.3

Example 1

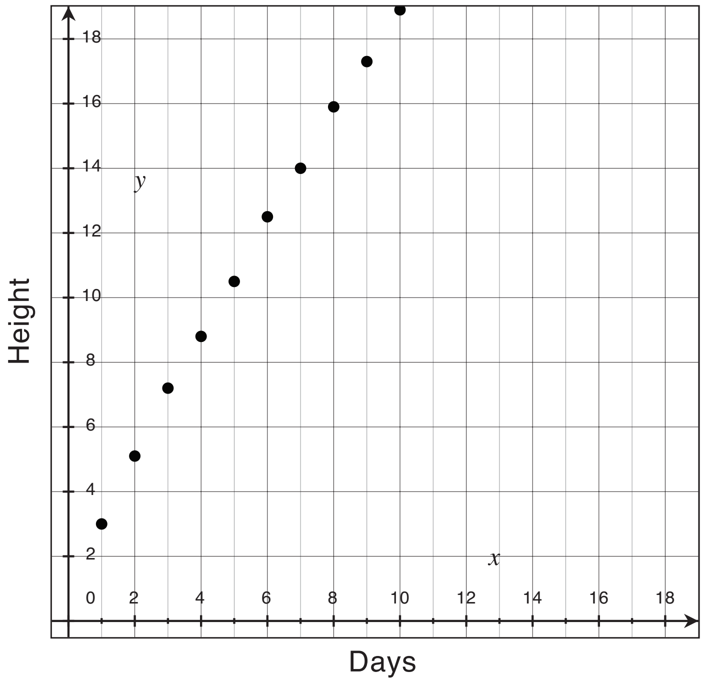
Pablo's science class is growing plants. He recorded the height of his plant each day for 10 days. The plant's height, in centimeters, over that time is listed in the table below.

Day	Height in centimeters
1	3
2	5.1
3	7.2
4	8.8
5	10.5
6	12.5
7	14
8	15.9
9	17.3
10	18.9

Pablo determines that the function $y = 1.73x + 1.87$ is a good fit for the data. How close is his estimate to the actual data? Approximately how much does the plant grow each day?

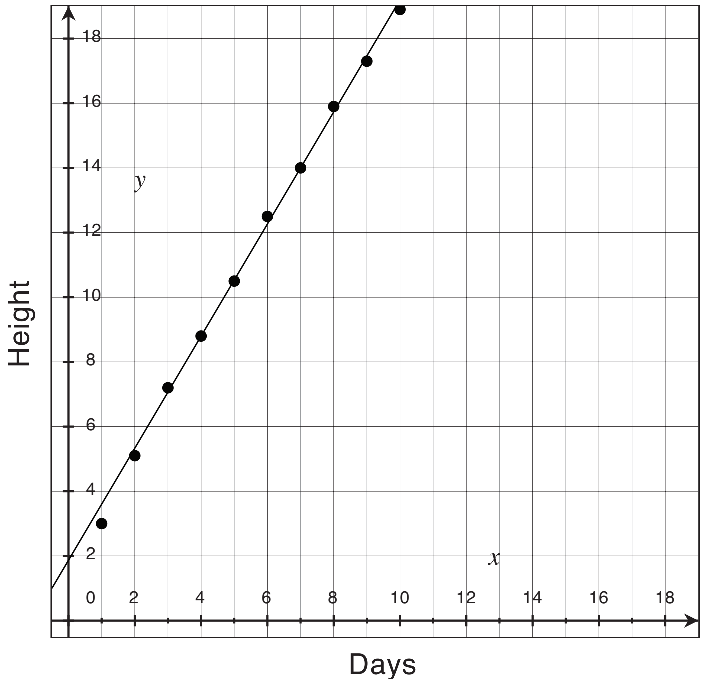
1. Create a scatter plot of the data.

   Let the *x*-axis represent days and the *y*-axis represent height in centimeters.



2. Draw the line of best fit through two of the data points.

   A good line of best fit will have some points below the line and some above the line. Use the graph to initially determine if the function is a good fit for the data.



3. Find the residuals for each data point.

The residual for each data point is the difference between the observed value and the estimated value using a line of best fit. Evaluate the equation of the line at each value of x.

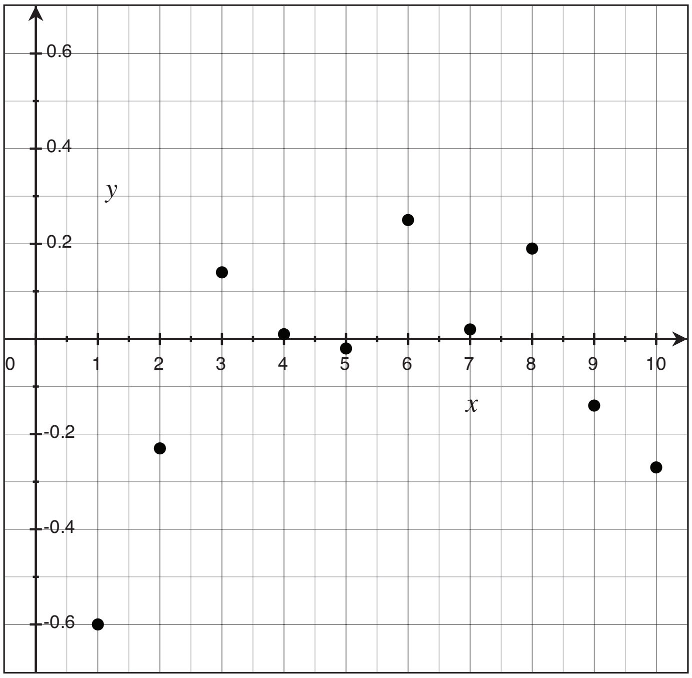
x	$y = 1.73x + 1.87$
1	$y = 1.73(1) + 1.87 = 3.6$
2	$y = 1.73(2) + 1.87 = 5.33$
3	$y = 1.73(3) + 1.87 = 7.06$
4	$y = 1.73(4) + 1.87 = 8.79$
5	$y = 1.73(5) + 1.87 = 10.52$
6	$y = 1.73(6) + 1.87 = 12.25$
7	$y = 1.73(7) + 1.87 = 13.98$
8	$y = 1.73(8) + 1.87 = 15.71$
9	$y = 1.73(9) + 1.87 = 17.44$
10	$y = 1.73(10) + 1.87 = 19.17$

Next, find the difference between each observed value and each calculated value for each value of x.

x	**Residual**
1	$3 - 3.6 = -0.6$
2	$5.1 - 5.33 = -0.23$
3	$7.2 - 7.06 = 0.14$
4	$8.8 - 8.79 = 0.01$
5	$10.5 - 10.52 = -0.02$
6	$12.5 - 12.25 = 0.25$
7	$14 - 13.98 = 0.02$
8	$15.9 - 15.71 = 0.19$
9	$17.3 - 17.44 = -0.14$
10	$18.9 - 19.17 = -0.27$

4. Plot the residuals on a residual plot.

   Plot the points (x, residual for x).



5. Describe the fit of the line based on the shape of the residual plot.

   The plot of the residuals appears to be random, with some negative and some positive values. This indicates that the line is a good line of fit.

6. Use the equation to estimate the centimeters grown each day.

   The change in the height per day is the centimeters grown each day. In the equation of the line, the slope is the change in height per day. The plant is growing approximately 1.73 centimeters each day.

Example 2

Lindsay created the table below showing the population of fruit flies over the last 10 weeks.

Week	Number of flies
1	50
2	78
3	98
4	122
5	153
6	191
7	238
8	298
9	373
10	466

She estimates that the population of fruit flies can be represented by the equation $y = 46x - 40$. Using residuals, determine if her representation is a good estimate.

1. Find the estimated population at each *x*-value.

   Evaluate the equation at each value of *x*.

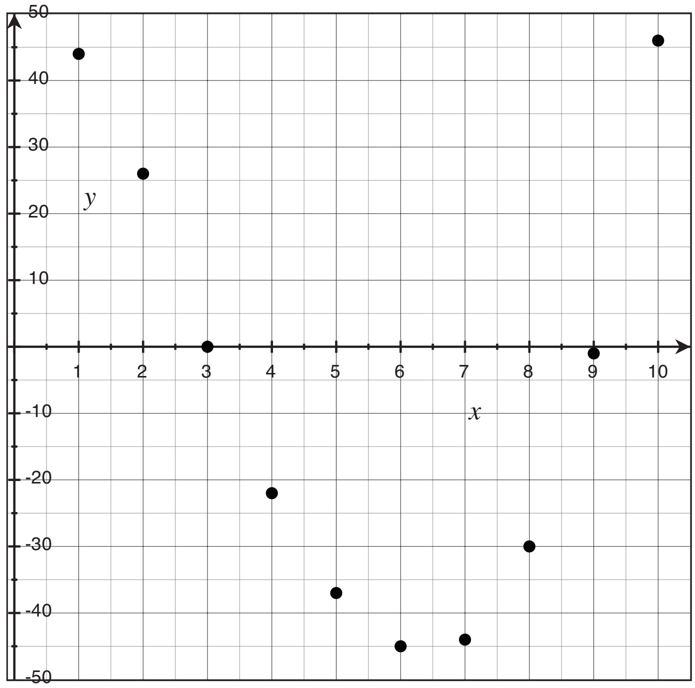
x	$y = 46x - 40$
1	$46(1) - 40 = 6$
2	$46(2) - 40 = 52$
3	$46(3) - 40 = 98$
4	$46(4) - 40 = 144$
5	$46(5) - 40 = 190$
6	$46(6) - 40 = 236$
7	$46(7) - 40 = 282$
8	$46(8) - 40 = 328$
9	$46(9) - 40 = 374$
10	$46(10) - 40 = 420$

2. Find the residuals by finding each difference between the observed population and estimated population.

x	Residual
1	50 – 6 = 44
2	78 – 52 = 26
3	98 – 98 = 0
4	122 – 144 = –22
5	153 – 190 = –37
6	191 – 236 = –45
7	238 – 282 = –44
8	298 – 328 = –30
9	373 – 374 = –1
10	466 – 420 = 46

3. Create a residual plot.

Plot the points (x, residual for x).



4. Analyze the residual plot to determine if the equation is a good estimate for the population.

   The residual plot has a U-shape. This indicates that a non-linear estimation would be a better fit for this data set.

   The shape of the residual plot indicates that the equation $y = 46x - 40$ is not a good estimate for this data set. ✓

Example 3

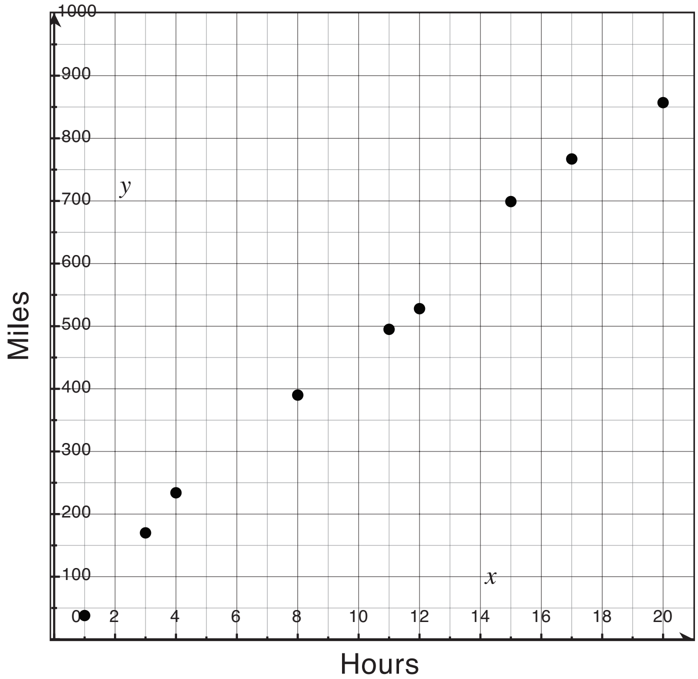
Anthony is traveling across the country by car. He keeps track of the hours he has driven and total miles he has traveled in the table below.

Hours	Miles
1	38
3	170
4	234
8	390
11	495
12	528
15	699
17	767
20	857

Anthony uses the equation $y = 42.64x + 42.12$ to estimate his total miles driven after any number of hours. Use a residual plot to determine how well the line fits the data. Approximately how many miles had Anthony driven after 13 hours?

1. Create a scatter plot of the data set.

   Let the x-axis represent hours and the y-axis represent miles.

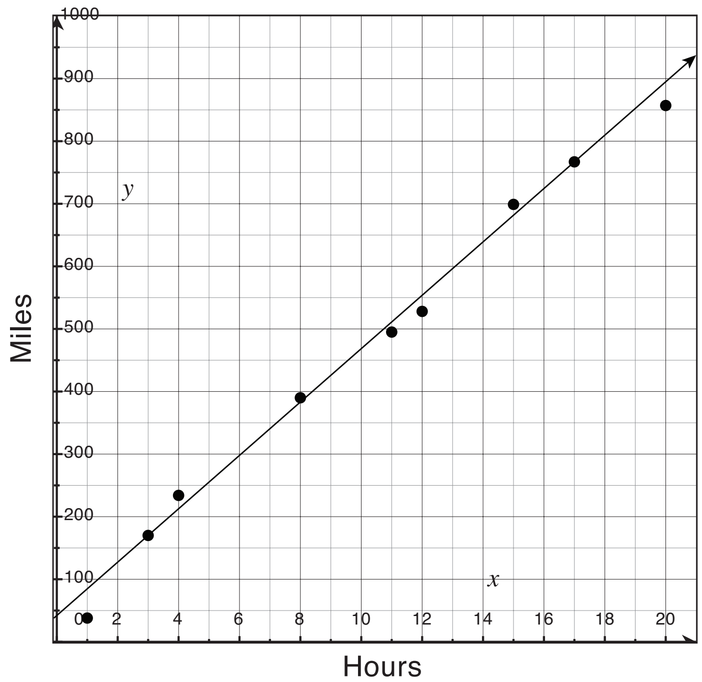


2. Plot the line of the equation Anthony used to estimate the total miles driven.

   To graph a linear equation, find two points on the line. Then draw a straight line through the two points. Two easy values of x to use are 0 and 1.

$y = 42.64(0) + 42.12 = 42.12$	Substitute 0 for x.
$y = 42.64(1) + 42.12 = 84.76$	Substitute 1 for x.

   Two points on the line are (0, 42.12) and (1, 84.76).



3. Find the residuals.

   Evaluate the line of best fit for each value of x.

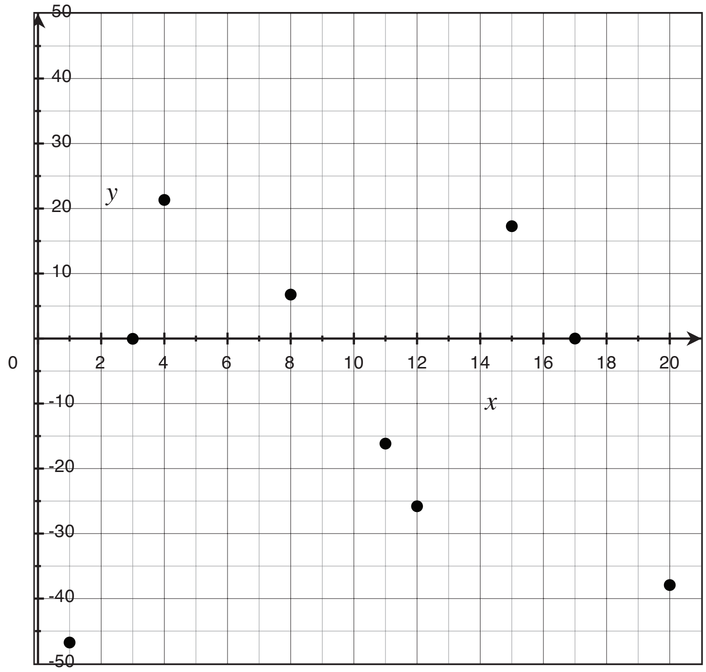
x	$y = 42.64x + 42.12$
1	42.64(1) + 42.12 = 84.76
3	42.64(3) + 42.12 = 170.04
4	42.64(4) + 42.12 = 212.68
8	42.64(8) + 42.12 = 383.24
11	42.64(11) + 42.12 = 511.16
12	42.64(12) + 42.12 = 553.8
15	42.64(15) + 42.12 = 681.72
17	42.64(17) + 42.12 = 767
20	42.64(20) + 42.12 = 894.92

4. Find the difference between each observed distance and estimated distance.

x	Residual
1	38 – 84.76 = –46.76
3	170 – 170.04 = –0.04
4	234 – 212.68 = 21.32
8	390 – 383.24 = 6.76
11	495 – 511.16 = –16.16
12	528 – 553.8 = –25.8
15	699 – 681.72 = 17.28
17	767 – 767 = 0
20	857 – 894.92 = –37.92

5. Create a residual plot.

   Plot the points (x, residual for x).



6. Determine if the linear function is a good estimate for the data.

   The residual plot has a random shape, indicating that the linear function is a good estimate for the data.

7. Use the equation to estimate the total miles driven when the time equals 13 hours.

   In the line of best fit, x = hours driven and y = total miles driven. Evaluate the function at $x = 13$ to estimate the total miles driven after 13 hours.

   $y = 42.64(13) + 42.12 = 596.44$ Substitute 13 for x.

   After 13 hours, Anthony had driven approximately 596 miles.