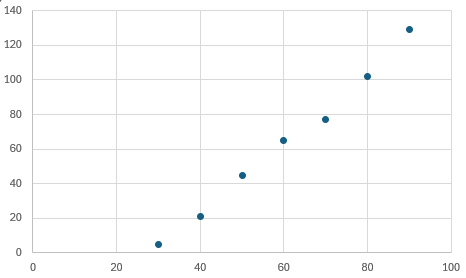
**Barron’s Let’s Review Regents – Algebra I**

# Chapter 12: Regression Curves

## 12.1 Line of Best Fit

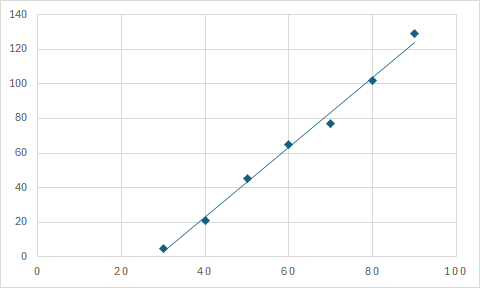
Two points can always be joined with a line. If there more than two points, there may not be one line that passes through all of them. If we need a linear equation for a line that comes close to all the points, there is a feature of the graphing calculator that will calculate the slope and y-intercept of this line.

|  |  |
| --- | --- |
| **Temperature** | **Number of people at the beach** |
| 30 | 5 |
| 40 | 21 |
| 50 | 45 |
| 60 | 65 |
| 70 | 77 |
| 80 | 102 |
| 90 | 129 |

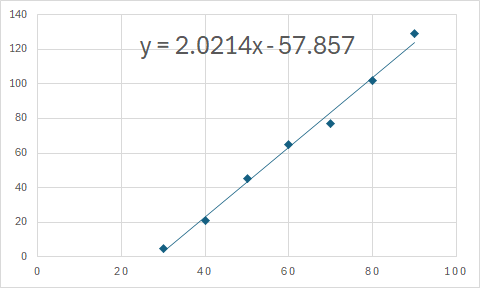


A line drawn through any of the two points will come close but will not pass through any of the other points.

Though there is not one line that passes through all the points, there is a *line of best fit* that will come close to all points. The equation for the line of best fit will be useful in calculating approximate coordinates for other points on the line.



Line of Best Fit with Equation



**Using the Line of Best Fit to Answer Questions About the Real-World Scenarios**

Approximately how many people will be at the beach when the temperature is 100 degrees?

Approximately 144 people will be there when it is 100 degrees.

At what temperature will there be 100 people on the beach?

At a temperature of approximately 78 degrees, there will be 100 people at the beach.

### Check Your Understanding of Section 12.1

1. Multiple-Choice
2. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 1 | 3 |
| 2 | 5 |
| 3 | 4 |
| 4 | 6 |
| 5 | 8 |

1. **y = 1.1x + 1.9**
2. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 1 | 7 |
| 2 | 6 |
| 3 | 6 |
| 4 | 5 |
| 5 | 4 |

**(4) y = -0.7x + 7.7**

1. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 1 | 1 |
| 2 | 3 |
| 3 | 8 |
| 4 | 7 |
| 5 | 9 |

1. **y = 2x - 0.4**
2. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 1 | 4 |
| 2 | 3 |
| 3 | 5 |
| 4 | 4 |
| 5 | 5 |

**(3) y = 0.3x + 3.3**

1. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 2 | 16 |
| 4 | 12 |
| 6 | 2 |
| 8 | 6 |
| 10 | 2 |

1. **y = -1.7x + 17.8**
2. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 10 | 33 |
| 20 | 20 |
| 30 | 10 |
| 40 | 14 |
| 50 | 6 |

**(2) y = -0.6x + 34.6**

1. Calculate the equation for the line of best fit for the following set of data in form. Round *m* and *b* to the nearest tenth.

|  |  |
| --- | --- |
| x | y |
| 10 | 250 |
| 25 | 310 |
| 37 | 450 |
| 46 | 560 |
| 59 | 820 |

**(3) y = 11.5x + 69.8**

1. What is the equation for the line of best fit for the points on this scatterplot?

|  |  |  |
| --- | --- | --- |
|  | x | y |
| A | 2 | 1 |
| B | 4 | 5 |
| C | 6 | 3 |
| D | 8 | 7 |
| E | 10 | 9 |

**(4) y = 0.9x - 0.4**

1. Of these four choices, which line appears to be the best fit for this scatterplot?  
   **(3)**
2. The equation is a line of best fit for which scatterplot?  
   **(1)**
3. *Show how you arrived at your answers*.
4. A bird-watching group tracks the number of birds they see and the temperature for six different locations. The data are collected in the table below. (a) Make a scatterplot of the data. (b) Calculate the line of best fit that relates birds (B) to temperature (T). Round values to the nearest hundredth. (c) Use your equation of the line of best fit to predict how many birds they would see if the temperature was 55 degrees.

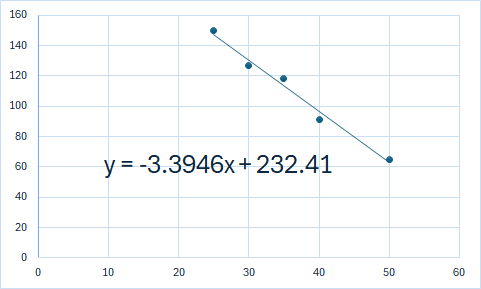
|  |  |
| --- | --- |
| **Temperature** | **Birds seen** |
| 40 | 30 |
| 50 | 41 |
| 60 | 72 |
| 70 | 91 |
| 80 | 94 |
| 90 | 89 |

**(a)** A graph with numbers and lines

AI-generated content may be incorrect.  
**(b)   
(c) or approximately 56 birds**

1. A car dealership keeps track of how many cars it sells at different prices. The data are collected in the table below. (a) Make a scatterplot of the data. (b) Calculate the line of best fit that relates cars sold (C) to price (P). Round values to the nearest hundredth. (c) Use your equation of the line of best fit to predict how many cars the dealership would sell at a cost of $31,000.

|  |  |
| --- | --- |
| Price (in thousands of dollars) | Cars sold |
| 25 | 150 |
| 30 | 127 |
| 35 | 118 |
| 40 | 91 |
| 50 | 65 |

**(a)**  
**(b)   
(c)   
Approximately 127 cars sold.**

1. How is it possible to find the equation for the line of best fit without a calculator?  
     
   Because the five points all lie on a line, the line of best fit is just the equation of the line through any of the two points. The equation is .
2. The two scatterplots below have all points the same, except the one on the left has the point (3,1) while the one on the right has the point (3,10). Will they have the same lines of best fit? If not, how will the line of best fit for the scatterplot on the left be different from the line of the best fit for the scatterplot on the right.  
     
   The scatterplot on the left with the point (3,1) will have a lower y-value on the y-intercept, and a greater slope, while the scatterplot on the right with the point (3,10) will have a higher y-value on the y-intercept and a slightly less increasing slope.

## 12.2 The Correlation Coefficient

The points on some scatterplots line up better than the points on other scatterplots. A measure of how well the points in a scatterplot resemble a straight line is called the correlation coefficient and is denoted by the letter r. The correlation coefficient is a number between -1 and +1. Correlation coefficients very close to -1 or +1 indicate that the points very nearly line up. Correlation coefficients close to 0 indicate that the points do not line up very well.

* A value close to +1 indicates a strong positive correlation.
* A value close to -1 indicates a strong negative correlation.
* A value close to 0 indicates a weak or no correlation.

**Calculating the Correlation Coefficient**

In Microsoft Excel:

**1. Prepare your data:**

Make sure your X and Y data are in separate columns in your Excel sheet.

**2.**Use the CORREL function:

* + Select an empty cell where you want to display the correlation coefficient.
  + Type =CORREL(.
  + Select the first dataset (the X values).
  + Type a comma (,).
  + Select the second dataset (the Y values).
  + Close the parentheses and press Enter.

**3. Interpret the result:**

The value returned by the CORREL function is the correlation coefficient, ranging from -1 to +1.

* + A value close to +1 indicates a strong positive correlation.
  + A value close to -1 indicates a strong negative correlation.
  + A value close to 0 indicates a weak or no correlation.

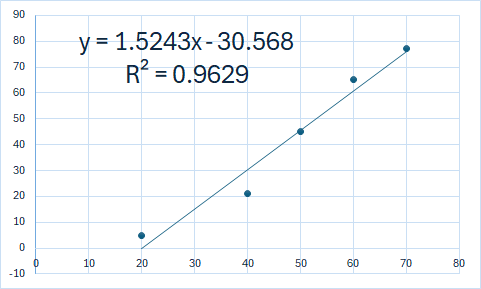
**4. Visualize the correlation:**

You can create a scatter plot using the X and Y data, and then add a trendline to visually see the correlation.

* + Select your data.
  + Go to the "Insert" tab and choose "Scatter".
  + Right-click on the data points and choose "Add Trendline".
  + Choose the "Linear" trendline type.
  + Check the "Display R-squared value on chart" box to see the coefficient of determination (R²).
  + The square root of R² is the correlation coefficient (R).

A screenshot of a computer

AI-generated content may be incorrect.



**Comparing Correlation Coefficients by Looking at the Scatterplots**

The data set from the temperature/people scenario had a correlation coefficient very close to 1. If the points on the scatterplot did not line up so nicely, the correlation coefficient would have been lower. Here are two data sets that have lower correlation coefficients.

The first set has a correlation coefficient of 0.94, and the second has a correlation coefficient of 0.84.

**Negative Correlation Coefficients**

If the line of best fit has a negative slope, the correlation coefficient will be negative. The more the points line up, the closer the correlation coefficient will be to -1, whereas the less they line up, the closer the correlation coefficient will be to 0.

### Check Your Understanding of Section 12.2

1. Multiple-Choice
2. What is the correlation coefficient (r), rounded to the nearest hundredth, for the line of best fit for the data in the table below?

|  |  |
| --- | --- |
| x | y |
| 3 | 10 |
| 6 | 13 |
| 9 | 27 |
| 12 | 38 |
| 15 | 40 |

**(1) 0.97**

1. What is the correlation coefficient (r) for the line of best fit for the data represented on the scatterplot below?  
   **(4) -6.7**
2. For which scatterplot is the correlation coefficient closest to 1?  
   **(2)**
3. Which scatterplot has a correlation coefficient closest to -1?  
   **(4)**
4. Which scatterplot has a correlation coefficient (r) closest to zero?  
   **(2)**
5. Of the four choices, which is closest to the correlation coefficient for this scatterplot?  
   **(1) 0.8**
6. Of the four choices, which is closest to the correlation coefficient for this scatterplot?  
   **(4) -0.7**
7. What is the correlation coefficient (r) for the line of best fit for the data represented on the scatterplot below?  
   **(3) 0.6864**
8. What could be the correlation coefficient for the line of best fit for a scatterplot that has only two points?  
   **(2) 1**
9. The scatterplot below has a line of best fit with equation and a correlation coefficient of 0.76. Which point added to the scatterplot, would increase the value of the correlation coefficient?  
   **(4) (6,7)**
10. *Show how you arrived at your answers*.
11. Which scatterplot has the higher correlation coefficient?  
      
    If you imagine a line running through the points of each scatterplot and mentally pictured the distance between the line and the points, you see a smaller distance between the points and the line on the left scatterplot and larger distance between the points and the line on the right scatterplot.  
      
    The smaller correlation coefficient resulting from points closer to the line would result in a coefficient value closer to 1 for the scatterplot on the left, and a coefficient value farther from one and closer to zero for the scatterplot on the right.
12. Ronald says that a correlation coefficient cannot be negative since completely random placed points have a correlation of 0 and the points cannot get any more randomly placed. Jeff says it is possible to have a negative correlation coefficient. Who is correct? Explain.  
      
    Jeff is correct. “If the line of best fit has a negative slope, the correlation coefficient will be negative. The more the points line up, the closer the correlation coefficient will be to -1, whereas the less they line up, the closer the correlation coefficient will be to 0.”
13. The seven points on the scatterplot all lie on the curve .  
      
    What can be determined about the correlation coefficient of the line of best fit for these seven points without using a calculator?  
      
    The correlation coefficient will be a positive value between 0 and 1 because the slope of the line of best fit will be positive, and it will not be equal to 1 because the curve is not a straight line.
14. Create a five-point scatterplot that has a line of best fit with the equation and a correlation coefficient (r) of +1.

|  |  |
| --- | --- |
| x | y |
| 1 | 4 |
| 2 | 5 |
| 3 | 6 |
| 4 | 7 |
| 5 | 8 |

A screenshot of a computer

AI-generated content may be incorrect.  
  
