

Name _____

Date _____

EXPLORATIONS

ACTIVITY 5

What Is a Linear Regression?

You *can* find the equation that best fits a set of data points!

Combined with the Texas Instrument Calculator Based Laboratory 2™ System (CBL 2™), the Voyage™ 200 PLT allows you to almost effortlessly fit a variety of different curves to data collected in real time. The resulting *regression equation* of the curve is then said to model the data. A regression equation model almost never models real data perfectly, but no matter what model you choose—linear, quadratic, cubic, quartic, exponential, power, or logarithmic, to name some of the regressions built into the Voyage 200 PLT—the mathematics for finding the best-fit regression equation is interesting in itself.

This exercise consists of two parts. In part one, you will create lists of data points, plot them on a graph, and then let the Voyage 200 PLT determine the regression equation of the line that best fits the data points. In part two, you will use the Voyage 200 PLT to understand exactly *how* the linear regression is determined.

Instructions

Part A—Finding a Linear Equation

Find a linear equation whose graph best fits the five points (1, 6), (2, 10), (3, 13), (4, 12), (5, 17).

1. Press **[ON]** to turn on the Voyage 200 PLT. To reset to the default settings, press **[2nd] [MEM] [F1]**, select 1:RAM, select 2:Default, and then press **[ENTER] [ENTER]**. Press **[♦]** **[CALC HOME]** to access the Home screen. Delete all values of one-character variables by pressing **[F6]**, selecting 1:Clear a–z, and pressing **[ENTER]**.

On the Home screen, clear the entry line by pressing **[CLEAR]** once or twice. Clear the History area above the entry line by pressing **[F1] 8**.

2. To begin, you need to make a list of the x -coordinates and store it under the variable name xs . At the entry line of the Home screen, press **[2nd] [1] 1 [2nd] [2] 2 [2nd] [3] 3 [2nd] [4] 4 [2nd] [5] 5 [2nd] [1] [STO] XS [ENTER]**.

Make a list of the y -coordinates, and store it under the variable name ys . Press

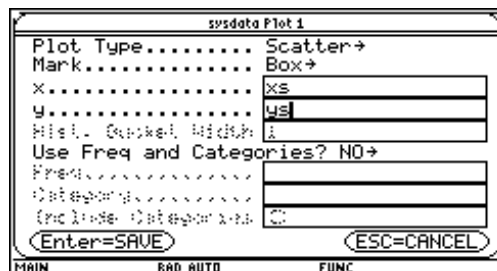
[2nd] [1] 6 [2nd] [2] 10 [2nd] [3] 13 [2nd] [4] 12 [2nd] [5] 17 [2nd] [1] [STO] YS [ENTER].

What Is a Linear Regression? (Continued)

- To set up a statistical data plot, press \blacktriangleright [Y=] to display the Y= Editor. If there are any functions or stat plots there, press $\boxed{F5}$ and select 1:All Off to deselect them.

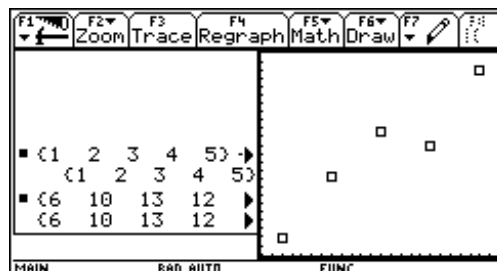
Press \blacktriangleleft to select the first empty plot above y1= and press \boxed{ENTER} . Plot Type and Mark are set to the desired defaults.

Scroll down with \blacktriangledown to the box x and type **XS**. Then scroll down to y and type **YS**. You will see the display at the right. Press \boxed{ENTER} \boxed{ENTER} to save.



- To split the screen, return to the Home screen by pressing \blacktriangleright [CALC HOME] and change the screen mode to split by pressing \boxed{MODE} $\boxed{F2}$ \blacktriangledown **3** \boxed{ENTER} . The screen is now split with the Home screen on the left and a graph screen on the right.

Graph the data points by pressing \blacktriangleright [GRAPH]. (The graph window is not set to display all the data points.) To fit the data points on the screen, press $\boxed{F2}$ and select 9:ZoomData. You should see the display at the right.



- Notice that the graph screen on the right has a black border around it. It is the *active* screen. Activate the Home screen on the left by pressing $\boxed{2nd}$ $\boxed{[⇄]}$. (You can “toggle” back and forth between screens in split screen mode by using $\boxed{2nd}$ $\boxed{[⇄]}$.)

What Is a Linear Regression? (Continued)

8. Guess values for m and k for a line to fit the data points. For an initial guess, choose $m = 2$ for the slope and $k = 2$ for the y -intercept. Then graph the line.

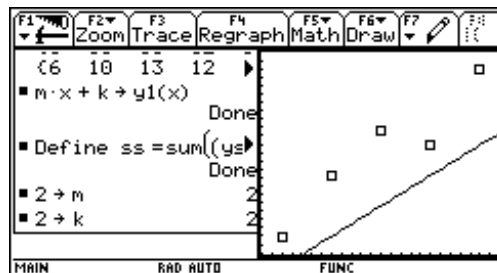
Press **ENTER** **CLEAR**

2 **STO** **M** **ENTER** **2** **STO** **K** **ENTER** **◆** **[GRAPH]**.

You should see the display on the right. It is not a particularly good fit.

Press **2nd** **[⇧]** to return to the Home screen. Now use the equation ss to measure the fit by pressing **SS** **ENTER**.

Because you assigned values to m and k , you should see the number 74.



9. With $m = 2$ and $k = 2$, the line $y = 2x + 2$ had a sum of squares of 74. Record those values in the first line of the table below.

From the graph of the line, you can tell that both the slope and y -intercept were a bit too low. Estimate new values of m and k for a line that better fits the data points by pressing

(new guess for m) **STO** **M** **ENTER**

(new guess for k) **STO** **K** **ENTER**

SS **ENTER**.

Record your values in the table below, and display the graph by pressing **◆** **[GRAPH]** to see the new line.

If your guess was an improvement, not only should the line look closer but the value of ss should be smaller.

Repeat this step several times until you are satisfied you have a line that appears to fit the data well. (Remember to toggle back to the Home screen by pressing **2nd** **[⇧]** after each new graph.)

$m =$	$k =$	$y = m \cdot x + k:$	$SS =$
$m =$	$k =$	$y = m \cdot x + k:$	$SS =$
$m =$	$k =$	$y = m \cdot x + k:$	$SS =$
$m =$	$k =$	$y = m \cdot x + k:$	$SS =$

What Is a Linear Regression? (Continued)

1. (continued)

Recall from algebra that the graph of $y = ax^2 + bx + c$ is a parabola opening upward when $a > 0$. Recall that the minimum value y of the function occurs at the vertex of the parabola.

Thus, a quadratic function in x has a minimum value when $x = \frac{-b}{2a}$.

2. To minimize the value of ss with respect to m , start by storing x into m .

Press **X** **[STO▶]** **M** **[ENTER]** **SS** **[ENTER]** to redisplay the expression.

Identify and record the expressions corresponding to a and b .

$a =$ _____ $b =$ _____.

Then,

$-b/2a =$ _____ is the value of m to minimize ss . Type these recorded expressions, and *store* them into a and b , respectively.

Delete the current value of m by pressing **[F4]**, selecting 4:DelVar, and pressing **M** **[ENTER]**.

Enter and store the first equation, which must hold true to minimize ss , by pressing **M** **[=]** **[(-)]** **B** **[=]** **[1]** **2** **[x]** **A** **[)]** **[STO▶]** **EQ1** **[ENTER]**.

3. To minimize the value of ss with respect to k , start by storing x into k , and then redisplay the expression.

Press **X** **[STO▶]** **K** **[ENTER]** **SS** **[ENTER]**.

Identify and record the expressions corresponding to a and b :

$a =$ _____ $b =$ _____

Then,

$-b/2a =$ _____ is the value of k to minimize ss . Type these recorded expressions, and *store* them into a and b , respectively.

What Is a Linear Regression? (Continued)

3. (continued)

Delete the current value of k by pressing $\boxed{F4}$, selecting 4:DelVar, and pressing \boxed{K} \boxed{ENTER} .

Enter and store the second equation, which must hold true to minimize ss , by pressing

\boxed{K} $\boxed{=}$ $\boxed{(-)}$ \boxed{B} $\boxed{=}$ $\boxed{1}$ $\boxed{2}$ $\boxed{\times}$ \boxed{A} $\boxed{)}$ $\boxed{STO\>}$ $\boxed{EQ2}$ \boxed{ENTER} .

4. You now need to solve the system of two equations in m and k . Because eq1 expresses m in terms of k , and eq2 expresses k in terms of m , solving the system is easy. Using a substitution method, it takes two steps:

(a) Solve the first equation for m , given that the second equation expresses k in terms of m .

Press

\boxed{SOLVE} $\boxed{1}$ $\boxed{EQ1}$ $\boxed{,}$ \boxed{M} $\boxed{)}$ $\boxed{2nd}$ $\boxed{1}$ $\boxed{EQ2}$ $\boxed{\blacklozenge}$ \boxed{ENTER} .

Record your answer below.

$m =$ _____

(b) Store the value for m .

(value of m) $\boxed{STO\>}$ \boxed{M} \boxed{ENTER}

Solve the second equation for k , given that the first equation expresses m in terms of k .

Press

\boxed{SOLVE} $\boxed{1}$ $\boxed{EQ1}$ $\boxed{,}$ \boxed{K} $\boxed{)}$ $\boxed{2nd}$ $\boxed{1}$ $\boxed{EQ2}$ $\boxed{\blacklozenge}$ \boxed{ENTER} .

Record your answer below.

$k =$ _____

Do these values agree with the regression coefficients found in step 6 of *Part A*?

What Is a Linear Regression? (Continued)

Extra Practice

Six students in a calculus class were given Voyage™ 200 Personal Learning Tools to help them improve their grades. The table below compares the number of hours each student practiced with the Voyage 200 PLT and his or her individual exam grade.

Assume that there is a linear correlation between the hours practicing with the Voyage 200 PLT and the grade. Use *Part A* of the previous activity as a model to find the values of m and k that produce the least sum of squares.

Check your solution by using the **LinReg**(command on the hours and the scores to calculate the linear regression directly.

hours worked with Voyage™ 200 PLT	score on calculus exam
4	62
9	99
2	48
5	64
7	85
8	92

What Is a Linear Regression? (Continued)

Extensions

The Exponential Regression

An exponential equation has the form $y = a * b^x$.

If you take the natural logarithm of both sides of this equation, the result is $\ln y = x * \ln b + \ln a$. This equation is linear; that is, $\ln(y)$ is a linear function of x . Therefore, you can perform a linear regression on the values of x and the values of $\ln(y)$.

Using the data from the table below, define lists for months and garbage. Define a third list as the natural logarithms of garbage. (Press **LN** **GARBAGE** **STO** **LNGARB** **ENTER**.)

Perform a linear regression on months and lngarb. (Press **LINREG** [space bar] **MONTHS** **▢** **LNGARB** **ENTER**.)

Exponentiate the regression coefficients. (Press **2nd** [e^x] **REGCOEF** **ENTER**). Compare this result with that obtained by performing an exponential regression directly on lists months and garbage. (Press **EXPREG** [space bar] **MONTHS** **▢** **GARBAGE** **ENTER** **SHOWSTAT** **ENTER**.)

You have calculated an exponential regression.

pounds of garbage in Sam Slobvik's dorm room after months away from home	
months	garbage
1	10
2	27
3	50
4	72

A Program for a Classroom Activity

See *Appendix A* for a listing of the program **BestFit()**. This program allows you to explore linear regression problems by interactively fitting a line to data points.

