Square It Up!
ID: 11409

Time Required
20 minutes

Activity Overview
In this activity, students will draw a scatter plot and a line. They will investigate the method of least squares by finding the residuals and the sum of the squares of the residuals. They will then test their line by using the built-in linear regression model.

Topic: Two-Variable Statistics
- *Scatter plots*
- *Lines of best fit*
- *Least squares regression*

Teacher Preparation and Notes
- *Have students clear all lists and functions before beginning the activity.*
- *This activity is intended to be student-centered. The worksheet is designed for students to work independently and then with a partner to answer a set of inquiry questions.*
- *Prior to this lesson, students should be able to construct a scatter plot.*
- ***To download the student worksheet, go to education.ti.com/exchange and enter "11409" in the keyword search box.***

Associated Materials
- *SquareItUp_Student.doc*

Suggested Related Activities
To download any activity listed, go to education.ti.com/exchange and enter the number in the keyword search box.
- *What is Linear Regression? (TI-84 Plus family) — 6194*
- *What's My Model? (TI-Nspire technology) — 8518*

One goal of this exploration is to see how the calculator finds the line of best fit, known as the linear regression model. It uses a technique called "Least Squares Regression." If possible, project or draw the scatter plot of the data on the front board of the classroom. Have a student draw her or his line of best fit. Draw line segments representing the residuals. Use these segments to draw squares. The sum of the areas of these squares is the quantity that we wish to minimize with the line of best fit.

Create a scatter plot

Students use the data on the worksheet to create a scatter plot of **Person** (*y* or L2) **vs. Year** (*x* or L1). They should describe the relationship of the variables as almost linear.

As students describe the data, they should note that there is a positive association between year and life expectancy.

Ask them to explain the meaning of the data. For example, in 1940 the life expectancy of a person was 63.6 years. Does that mean that all of those people born in 1940 have passed away?
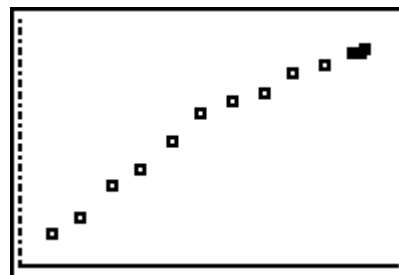
Draw a line that fits the data

The line that is given on the worksheet is below most of the data points. Students should find a line that fits the data better than this one.

Each student will have a different line, but each line should follow the "line of best fit" properties.

1. There should be an equal number of points above and below the line.

2. The points should be distributed throughout the line, there should not be a concentration of points above (or below) the line in one area.

If necessary, remind students how to calculate the equation of a line given the coordinates of two points on it.
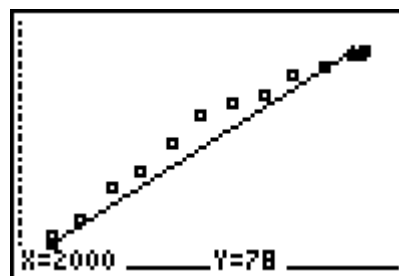
Compute the sum of the squares of the residuals

Encourage your students to try additional lines of "better fit" in order to minimize the sum of the squares of the residuals.

A good sum will be somewhere between 41 and 50.

Find the line of best fit

The equation of the line of best fit is shown below on the left. The sum of the squares of the residuals is shown on the right.

```
LinReg
 y=ax+b
 a=.2661288633
 b=⁻454.3637541

■
```

```
a=.2661288633
b=⁻454.3637541



sum(L₄)
       41.09625633
```

If using MathPrint™ OS:

When entering the command for the linear regression, students can enter Y2 by either (1) press $\boxed{\text{VARS}}$ $\boxed{\blacktriangleright}$ $\boxed{\text{ENTER}}$ and select Y2 or (2) press $\boxed{\text{ALPHA}}$ [F4] and select Y2.

```
                    Y1  Y6
                    Y2  Y7
                    Y3  Y8
                    Y4  Y9
                    Y5  Y0
FRAC FUNC MTRX YVAR
```

If using MathPrint™ OS:

When students choose **LinReg(ax+b)** from the CALC menu, a screen similar to that shown at the right will be displayed. Students should enter the Xlist, Ylist, a location (e.g., **Y2)** to store the regression equation, and then select Calculate.

```
  LinReg(ax+b)
Xlist:L₁
Ylist:L₂
FreqList:
Store RegEQ:
Calculate
```

Homework

Students can analyze the two additional data sets and fill in the table on the student worksheet. Have students determine what is similar and different between the three scatter plots and regression lines.

	Women	*Men*
Your equation	Answers will vary	Answers will vary
Sum of squares	Answers will vary	Answers will vary
Calculator equation	$y = 0.280241x - 479.52$	$y = 0.248491x - 422.244$
Sum of squares	66.3362	30.2809

Extensions

1. Can one use the linear regression model found to predict the life expectancy in 2010? In 2020?

2. Can one predict what year the life expectancy will be 92? 105?

3. What years and life expectancies make sense in the model?

4. Find the means of Year and Person (Answer: 1964 and 68.3133). Is that point on the graph of the Least Squares Regression Equation? (Answer: Yes)