Square It Up!

Name _____

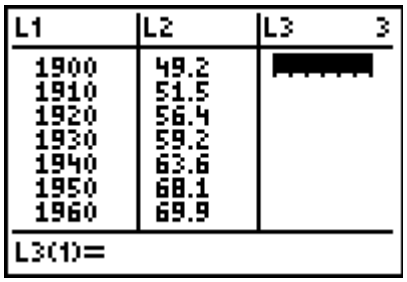
Class _____

Create a scatter plot

The table on the right shows the life expectancies from 1900 to 2004 for men, women, and people (men and women combined). Investigate the relationship between the year and the life expectancy for a person living during that year by creating a scatter plot.

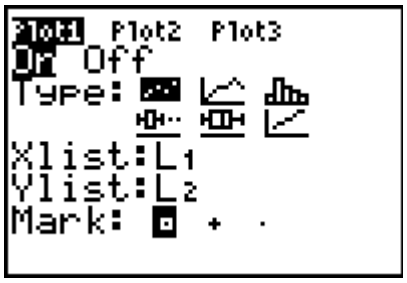
Year	Person	Men	Women
1900	49.2	47.9	50.7
1910	51.5	49.9	53.2
1920	56.4	55.5	57.4
1930	59.2	57.7	60.9
1940	63.6	61.6	65.9
1950	68.1	65.5	71
1960	69.9	66.8	73.2
1970	70.8	67	74.6
1980	73.9	70.1	77.6
1990	75.4	71.8	78.8
2000	77	74.3	79.7
2001	77.2	74.4	79.8
2002	77.3	74.5	79.9
2003	77.4	74.7	80
2004	77.8	75.2	80.4

Step 1: Press [STAT] [ENTER]. Enter the data for **Year** in list L1 and the data for **Person** in list L2.



Step 2: Create the scatter plot by pressing [2nd] [STAT PLOT] and set up **Plot1** as shown on the right.

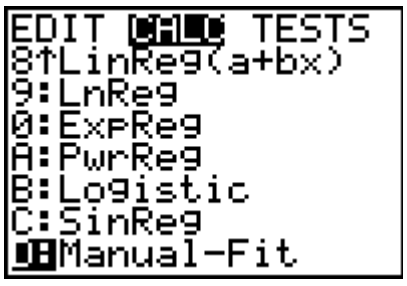
Press [ZOOM] and select **ZoomStat** to adjust the window settings appropriately.



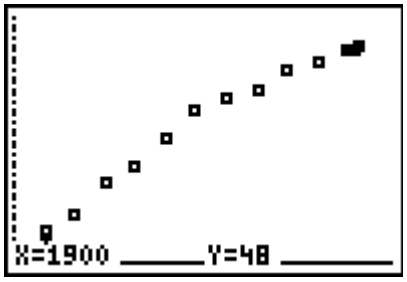
- Describe the relationship between the two variables.

Draw a line that fits the data

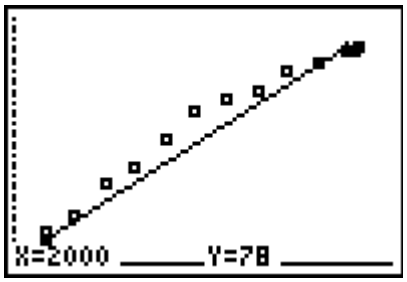
Step 3: Draw a line that fits the data as well as possible. Press [STAT] [▸] to access the CALC menu and then choose **Manual-Fit**.



Step 4: Move the cursor to one side of the scatter plot and anchor the point by pressing [ENTER].



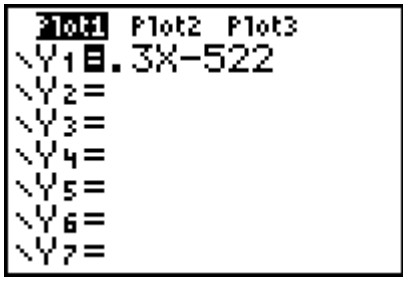
Step 5: Move the cursor to the other side of the scatter plot and press [ENTER] to anchor the right endpoint. The equation of your line will appear in the top left corner.



Step 6: Modify the graph until you have a line that fits the data better than the line shown to the right. This line will be your line of "better fit."

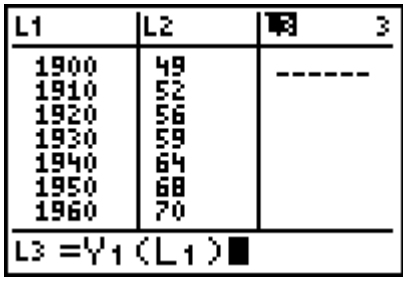
Step 7: Press [2nd] [QUIT] to finish the Manual Fit function. The equation of your line is saved in Y_1.

- What is the equation of your line?



Compute the sum of the squares of the residuals

Step 8: Compare the coordinates of the points on your line with those of the original data by entering **$Y_1(L_1)$** in the top of list L_3. Press [ENTER].



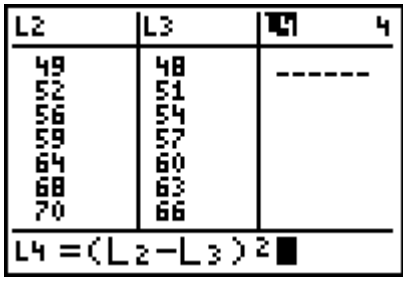
Residuals are the differences between the actual *y*-values of the data set and the corresponding *y*-values predicted by your line of "better fit."

Residual = Actual value – Predicted value

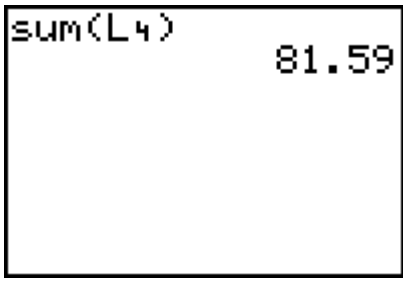
- If a data point is *above* the line, is the value of the residual positive or negative?

- If a data point is *below* the line, is the value of the residual positive or negative?

Step 9: Find the square of these residuals. Move to the top of list L4 and enter **$(L_2 - L_3)^2$**. Press [ENTER].



Step 10: Return to the Home screen. Find the sum of the squares of the residuals. Press [2nd] [LIST] and under the MATH menu select **sum(**. Then enter **L4)**. Press [ENTER].



- What is your sum?

Did you get a value lower than the one shown? If so, then you found a line of "better fit."

- Compare the sum of the squares of your residuals with those of your classmates. What was the lowest sum?

- How many points are above their line? Below their line?

- Which data points do you think increase the sum of the squares the most?

- How do the values in list L4 for these data points compare to the others?

- Where are these points located in relation to the line?

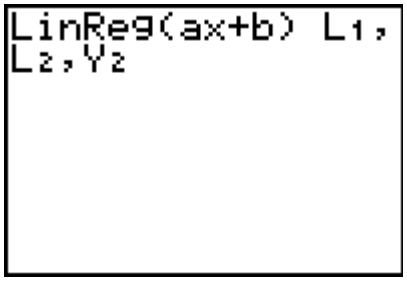
- What do you notice about the distribution of the data points around the line (i.e., above vs. below the line, or equal spacing vs. clusters)

- To have the smallest sum, how do the points need to be distributed?

Find the line of best fit

Step 11: There is a line of best fit. That line is called the Least Squares Regression Line, often written as LSRL.

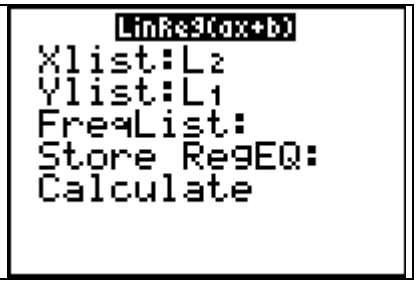
Press [STAT] ▸ and select **LinReg(ax+b)** from the CALC menu. Then enter **L1, L2, Y2**. Press [ENTER].



This command computes the equation of the Least Squares Regression Line and stores it in Y2. Graph this line, along with your line of "better fit," and the scatter plot.

If using MathPrint™ OS:

Choose **LinReg(ax+b)** from the CALC menu. A screen similar to that shown at the right will be displayed. Enter the Xlist, Ylist, a location (**Y2)** to store the regression equation, and then select Calculate.



- How does the graph of the LSRL compare with the graph of your line of "better fit?"

Step 12: Repeat Steps 8–10 to find the sum of the squares of the residuals for the LSRL which is stored in Y2.

- What is the value of the sum of squares?

- How well did you do? Does your sum of squares compare favorably to the sum of squares for the calculator-generated LSRL equation?

Homework

Analyze the two additional sets of data in the spreadsheet by using **Manual-Fit** to draw a line that fits the data. Compute the sum of the squares for your equation. Compare your answers each time to the Least Squares Regression Model and the sum of the squares for the regression equation.

	Women	Men
Your equation		
Sum of squares		
Calculator equation		
Sum of squares		