Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Period: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 **TV Watching vs. Physical Activity**

**Overview & Instructions**



Do people who watch a lot of TV not do much physical activity?

Check out this relationship in your classroom.

1. Write down the number of hours a week you watch TV and the number of hours you do physical activity on the first line of the given chart.
	* Make sure that you give each classmate the same information when asked.

2. Next, go around with this page and write down two pieces of information from each classmate and ask:

* How many hours of television each student watches approximately per week?
* How many and how many hours of physical activity do you do?

3. Next, make a graph with hours of TV watched on the x-axis and hours of physical activity on the y-axis using the next page attached in this packet.

4. Then, find the coefficient correlation of your data in order to find the “line of best fit” for your data using a graphing calculator.

5. Finally, you will find the regression model of the data using a graphing calculator.

6. To end the class we will discuss as a class our conclusions from this lesson.

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**TV Watching vs. Physical Activity Data Sheet**

|  |  |  |
| --- | --- | --- |
| Subject | TV Watching (hours) | Physical Activity (hours) |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Regression Worksheet**

1. Make a scatter plot of the data using hours of TV watched as the independent variable (x-axis) and physical activity as the dependent variable (y-axis). Please do not forget to label your axis!



1. Before doing any mathematical calculations to find the correlation, based on the data (plotted points) what is the correlation? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
2. Using your best mathematical reasoning, why do you think correlation is important? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Regression Worksheet Cont.**

1. Following teacher’s instructions, or the Student References (attached below: pgs. 5-6) please insert your data into your graphing calculator.
2. Following teacher’s instructions, or the Student References (attached below: pgs. 7-9) please create a scatter plot using your graphing calculator. Then, find two classmates to compare your scatter plot with. If your graphs do not coincide, please ask teacher for assistance.
3. Following teacher’s instructions, or the Student References (attached below: pgs. 10-11) please choose at least two regression models using your graphing calculator.
4. Is there any correlation in our data is it a strong correlation? Please explain why or why not using complete sentences and mathematical terminology. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Get together with a partner and brainstorm on problems (8-12).

1. on which of the (four different) regression models you chose is the “best fit” for this regression model and make sure to use mathematical terms to prove your argument in the space provided below:
	* Hint: Remember in order to prove that your regression model is correct, make sure that your correlation coefficient is high.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. Then, individually graph the line of best fit you both decided on onto your plotted graph (on previous page). Use a ruler if necessary. Be as precise as possible.
2. Write the regression model equation: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
3. What is the y-intercept of the regression model? Does it have any significance to our data? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
4. What is the slope of the regression model? Does it have any significance to our data?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. Brainstorm: Suppose a person watches 50 hours of television per week. How much physical activity is expected that he or she does? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**Student Reference:** **Steps on How to Insert Data Using a Graphing Calculator**

**The TI-83 Calculator can store data into lists.  This is necessary for a number of functions, including Scatterplots and Linear Regression.**

Example: Insert and Name a list "Age" using the following list of data

|  |
| --- |
| **X** |
| 28 |
| 48 |
| 53 |
| 25 |
| 38 |
| 23 |
| 49 |
| 32 |

|  |  |
| --- | --- |
| **Step 1:** Press STAT.  This is the EDIT menu.    | linreg8 |

|  |  |
| --- | --- |
| **Step 2:** Press ENTER to edit lists.   | insert1 |
| **Step 3:** Use the blue up arrow key to highlight the list where the new list will be inserted.     | insert2 |

|  |  |
| --- | --- |
| **Step 4:** Press the yellow 2ND key, then INS above DEL.  This will insert a new list to the *left* of the list highlighted in the previous step.  Note the bottom left hand side of the screen has "Name=".        | insert3  |
| **Step 5:** Enter the name of the list.   *In this example, enter AGES.*Use the appropriate letter keys in green above the buttons on the calculator.  Note: There is no need to press the green ALPHA key.  The calculator is automatically set to accept letters.*In this example...**A is located above the MATH keyG is located above the TAN keyE is located above the SIN keyS is located above the LN key* | insert4 |
| **Step 6:** Press ENTER.  The list name should appear at both the top of the column and at the bottom of the screen.   | insert5 |

You can now enter data into this list as you would any other.

Teacher Note: Please note if you use 1-VAR STATS or Linear Regression for a named lists -- that is, lists other than the default L1, L2, etc. -- you will need to enter that list's name using the green ALPHA key and the green letters in lieu of the default lists.

**Student Reference:** **Steps on How to Create a Scatter Plot Using a Graphing Calculator**

### Step 0. Setup

|  |  |
| --- | --- |
| Set floating point mode, if you haven’t already. | [MODE] [▼] [ENTER] |
| Go to the home screen | [2nd MODE *makes* QUIT] [CLEAR] |
| Turn on diagnostics with the [DiagnosticOn] command.  | [2nd 0 *makes* CATALOG] [x-1] Don’t press the [ALPHA] key, because the CATALOG command has already put the calculator in alpha mode.  Scroll down to DiagnosticOn and press [ENTER] twice. |

The calculator will remember these settings when you turn it off: next time you can start with Step 1.

### Step 1. Make the Scatter Plot

Before you even run a regression, you should **first plot the points and see** whether they seem to lie along a **straight line**. If the distribution is obviously not a straight line, don’t do a linear regression. (Some other form of regression might still be appropriate, but that is outside the scope of this course.)

Let’s use this example from [Sullivan (2011, 179)](http://www.tc3.edu/instruct/sbrown/swt/sources.htm#so_Sullivan2011) [see “Sources Used” at end of book]: the distance a golf ball travels versus the speed with which the club head hit it.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Club-head speed, mph (x)** | 100 | 102 | 103 | 101 | 105 | 100 | 99 | 105 |
| **Distance, yards (y)** | 257 | 264 | 274 | 266 | 277 | 263 | 258 | 275 |

|  |  |
| --- | --- |
| Turn off other plots. | [Y=] Cursor to each highlighted = sign or Plot number and press [ENTER] to deactivate. |
| Set the format screen. | TI-83/84 format screenPress [2nd ZOOM *makes* FORMAT]. Just select everything in the left column.  |
| Enter the numbers in two statistics lists. | [STAT] [1] selects the list-edit screen.  Cursor onto the label L1 at top of first column, then [CLEAR] [ENTER] erases the list. Enter the x values.  Cursor onto the label L2 at top of second column, then [CLEAR] [ENTER] erases the list. Enter the y values. |
| Set up the scatterplot. | [2nd Y= *makes* STAT PLOT] [1] [ENTER] turns Plot 1 on. |
|   | [▼] [ENTER] selects scatterplot. |
|   | [▼] [2nd 1 *makes* L1] ties list 1 to the x axis. |
|   | [▼] [2nd 2 *makes* L2] ties list 2 to the y axis. |
| Plot the points. I have the grid turned on in some of these pictures, but [earlier](http://www.tc3.edu/instruct/sbrown/swt/chap04.htm#c04_FormatScreen) I told you to turn it off. That’s simplest. If you want the grid, you can turn it on, but then you’ll have to adjust the grid spacing for almost every plot. To adjust grid spacing, press [WINDOW], set Xscl and Yscl to appropriate values for your data, and press [GRAPH] to see the result. | TI-83/84 scatterplot[ZOOM] [9] automatically adjusts the window frame to fit the data.  |
| Check your data entry by tracing the points. | TI-83/84 scatterplot with trace for checking data entry[TRACE] shows you the first (x,y) pair, and then [►] shows you the others. They’re shown in the order you entered them, not necessarily from left to right.  |

A scatterplot on paper needs labels (numbers) and titles on both axes; the x and y axes typically won’t start at 0. Here’s the plot for this data set. (The horizontal lines aren’t needed when you plot on graph paper.)

When the same (x,y) pair occurs multiple times, plot the extra ones slightly offset. This is called **jitter**. In the example at the right, the point (6,6) occurs twice.

If the data points don’t seem to follow a straight line reasonably well, **STOP!** Your calculator will obey you if you tell it to perform a linear regression, but if the points don’t actually fit a straight line then it’s a case of “garbage in, garbage out.”

For instance, consider this example from [DeVeaux, Velleman, Bock (2009, 179)](http://www.tc3.edu/instruct/sbrown/swt/sources.htm#so_DeVeaux2009) [see “Sources Used” at end of book]. This is a table of recommended f/stops for various shutter speeds for a digital camera:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Shutter speed (x)** | 1/1000 | 1/500 | 1/250 | 1/125 | 1/60 | 1/30 | 1/15 | 1/8 |
| **f/stop (y)** | 2.8 | 4 | 5.6 | 8 | 11 | 16 | 22 | 32 |

If you try plotting these numbers yourself, enter the shutter speeds as fractions for accuracy: don’t convert them to decimals yourself. The calculator will show you only a few decimal places, but it maintains much greater precision internally.

You can see from the plot at right that these data don’t fit a straight line. There is a distinct bend near the left. When you have anything with a curve or bend, linear regression is wrong. You can try other forms of regression in your calculator’s menu, or you can transform the data as described in [DeVeaux, Velleman, Bock (2009, ch 10)](http://www.tc3.edu/instruct/sbrown/swt/sources.htm#so_DeVeaux2009) [see “Sources Used” at end of book] and other textbooks.

Reference: <http://www.tc3.edu/instruct/sbrown/swt/chap04.htm#c04_regress_root>

**Student Reference:** **Steps on How to Perform the Regression Using a Graphing Calculator**

### Step 2. Perform the Regression

|  |  |
| --- | --- |
| Set up to calculate statistics. | [STAT] [►] [4] pastes LinReg(ax+b) to the home screen. |
|   | [2nd 1 *makes* L1] [,] [2nd 2 *makes* L2] defines L1 as x values and L2 as y values.  If you have the “wizard’ interface, leave FreqList blank, or press [DEL] if something is already filled in. |
| Set up to store regression equation. | [,] [VARS] [►] [1] [1] pastes Y1 into the LinReg command. |
| Show your work! Write down the whole command — LinReg(ax+b) L1,L2,Y1 in this case, not just LinReg or LinReg(ax+b).  | Press [ENTER]. The calculator shows correlation and regression statistics and pastes the regression equation into Y1. |

Your input screen should look like this, for the “wizard” and non-wizard interfaces:

      

Write down the slope a, the y intercept b, the coefficient of determination R², and the correlation coefficient r. (A decent rule of thumb is four decimal places for slope and intercept, and two for r and R².)

a = 3.1661, b = −55.7966

R² = 0.88, r = 0.94

Reference: <http://www.tc3.edu/instruct/sbrown/swt/chap04.htm#c04_regress_root>

**Student Reference:** **How to Display the Regression Line**

### Step 3. Display the Regression Line

|  |  |
| --- | --- |
| Show line with original data points. | TI-83/84 plotted points and regression line[GRAPH] |

What is this line, exactly? It’s the one unique line that fits the plotted points best. But what does “best” mean?



The same four points on left and right. The vertical distance from each measured data point to the line, y−ŷ, is called the residual for that x value. The line on the right is better because the residuals are smaller.
source: [Dabes and Janik (1999, 179)](http://www.tc3.edu/instruct/sbrown/swt/sources.htm#so_Dabes1999) [see “Sources Used” at end of book]

For each plotted point, there is a **residual** equal to y−ŷ, the difference between the actual measured y for that x and the value predicted by the line. Residuals are positive if the data point is above the line, or negative if the data point is below the line.

You can think of the residuals as measures of how bad the line is at prediction, so you want them small. For any possible line, there’s a “total badness” equal to taking all the residuals, squaring them, and adding them up. The **least squares regression line** means the line that is best because it has less of this “total badness” than any other possible line. Obviously you’re not going to try different lines and make those calculations, because the formulas built into your calculator guarantee that there’s one best line and this is it.

[Carl Friedrich Gauss](http://www.tc3.edu/instruct/sbrown/swt/bignames.htm#bign_Gauss) developed the method of least squares in a paper published in 1809.

Reference: <http://www.tc3.edu/instruct/sbrown/swt/chap04.htm#c04_regress_root>