**Finding the Line of Best Fit with Technology**

We have learned how to fit a trend line to a scatter plot and find an equation of the trend line. In this activity we will use the calculator to perform both tasks. The trend line you will find with the calculator is called the **regression line** or **line of best fit**.

The table below, from Activity 5.2.2, gives the height (in inches) and weight (in pounds) of some of the NBA’s greatest players.

|  |  |  |
| --- | --- | --- |
| Player | Height | Weight |
| Kareem Abdul-Jabbar | 86 | 266 |
| Larry Bird | 80 | 220 |
| Wilt Chamberlain | 85 | 275 |
| Patrick Ewing | 84 | 255 |
| Magic Johnson | 83 | 255 |
| Michael Jordan | 78 | 215 |
| Scottie Pippen | 79 | 228 |
| Isaiah Thomas | 73 | 182 |

1. Enter the height data in L1 and the weight data in L2. Do not sort the data because the height and weight of each player need to stay together. Your lists should look like figure 1 below.
2. You can graph these points by turning on a STAT PLOT, choosing a scatter plot, and using L1 and L2 for the lists. See figure 2 below.
3. To see your scatter plot, you first need to set your window. We put the height data in L1, which the calculator is using as the *x*-coordinates, so *x* needs to go from 73 to 86. Set the *x*-scale to 2.

Since *y* contains the weights, it needs to go from 182 to 266.

Set the *y*-scale to 10. It is always a good idea to set

your window a bit bigger to see all the points.

See figure 3 below.

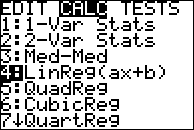
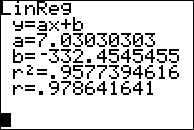
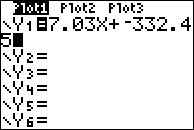
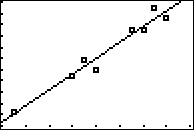
1. Hit the GRAPH button to see the scatter plot. See figure 4 below.

Figure 1 Figure 2 Figure 3 Figure 4

|  |
| --- |
| [image][image][image][image] |

1. Your calculator can find the equation of the line of best fit through a process called **linear regression**. This is a statistical calculation, so the command is found in the STAT ► CALC menu ▼ 4:LinReg(ax+b). Note that the calculator uses *a* for the slope and *b* for the *y*-intercept. Hit ENTER to bring the command to the home screen and ENTER again to run the LinReg(ax+b) command. This command is shown in figure 5 on the next page.

Figure 5 Figure 6 Figure 7 Figure 8



The output of the LinReg(ax+b) command is shown in Figure 6. *a* ≈ 7.03 is the slope of the regression line and *b* ≈ –332.45 is the *y*-intercept of the regression line. Using these values, the equation of the regression line is . To graph the regression line, type the equation into the Y= menu as shown in figure 7. Hit Graph. The regression line will appear in the same window as the scatter plot as shown in figure 8.

The statistics *r*2 and *r* in the LinReg(ax+b) measure the **correlation** between the two variables.

*r* is called the **correlation coefficient**. It measures the strength and direction of the **linear relationship** between two variables. The sign (positive or negative) of *r* indicates the direction of the linear relationship. If *r* is positive, the slope of the regression line is positive, and if *r* is negative, the slope of the regression line is negative. The value of *r* measures the strength of the linear relationship. If *r* is close to 1 or -1, the data has a strong linear relationship.

Scatterplots are shown below with their corresponding values of *r*.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Perfect Positive Correlation |  | High  Positive Correlation |  | Low  Positive Correlation |  | No Correlation |  | Low Negative Correlation |  | High  Negative Correlation |  | Perfect Negative  Correlation |
|  | | | | | | | | | | | | | |

1. Circle the word(s) that describe the correlation between basketball players’ height and weight. Explain your choice.

strong weak none

positive negative

For each data set below, state an appropriate window, make a scatter plot, and find and graph the line of best fit using your calculator. Note the correlation of the two variables in each example. Show your teacher each completed screen.

1. The table below from Activity 5.2.2 gives the city and highway miles per gallon (MPG) used by some of 2010’s most fuel-efficient cars. Enter city MPG in L1 and highway MPG in L2.

|  |  |  |
| --- | --- | --- |
| Car | City MPG | Highway MPG |
| Toyota Prius | 51 | 48 |
| Honda Civic Hybrid | 40 | 45 |
| Honda Insight | 40 | 43 |
| SmartCar | 33 | 41 |
| Volkswagon Jetta | 30 | 41 |
| Toyota Yaris | 29 | 36 |
| MINI Cooper | 28 | 37 |
| Kia Rio | 28 | 34 |
| Hyundai Elantra | 26 | 35 |
| Honda Accord | 22 | 31 |

1. Window: *X*min \_\_\_\_\_ *X*max \_\_\_\_\_ *X*scl \_\_\_\_\_

*Y*min \_\_\_\_\_ *Y*max \_\_\_\_\_ *Y*scl \_\_\_\_\_

1. Equation of the Line of Best Fit:

*y* = \_\_\_\_\_\_\_\_\_\_\_ *x*  + \_\_\_\_\_\_\_\_\_\_\_\_

1. What is the value of *r*?
2. Circle the word(s) that describe the correlation between city MPG and highway MPG. Explain your choice.

strong weak none

positive negative

1. What is the predicted highway MPG for a car that gets 38 MPG in the city?
2. The table below shows the number of viewers (in millions) and the rank for a sample of shows from the 2007-2008 season. Enter millions of viewers in L1 and rank in L2.

|  |  |  |
| --- | --- | --- |
| Name of Show | Viewers (millions) | Rank |
| How I met your Mother | 8.21 | 70 |
| New Amsterdam | 8.85 | 60 |
| Deal or no Deal | 9.72 | 50 |
| Sunday Night NFL Pre-Kick Off | 10.64 | 40 |
| Law and Order: SVU | 11.33 | 30 |
| Without a Trace | 13.13 | 20 |
| Grey’s Anatomy | 15.92 | 10 |

1. Window: *X*min \_\_\_\_\_ *X*max \_\_\_\_\_ *X*scl \_\_\_\_\_

*Y*min \_\_\_\_\_ *Y*max \_\_\_\_\_ *Y*scl \_\_\_\_\_

1. Equation of the Line of Best Fit:

*y* = \_\_\_\_\_\_\_\_\_\_\_ *x*  + \_\_\_\_\_\_\_\_\_\_\_\_

1. What is the value of *r*?
2. Circle the word(s) that describe the correlation between viewers (in millions) and rank. Explain your choice.

strong weak none

positive negative

1. What is the predicted rank of a show with 5 million viewers?