Linear Regressions
Getting started with Making a Scatter Plot on your Calculator
Group Activity 1
STEM Project Week #3

Regression analysis is a method used to create a mathematical model to show the relationship between two quantities. These quantities are usually represented by variables. One of these variables is the dependent variable and the other is the independent variable. (Remember that an independent variable represents the value being changed, usually \( x \), and the dependent variable is the variable that changes, usually \( y \). \( y \) depends on \( x \).)

A regression equation is an equation that "best fits" a given set of data. It best describes the relationship between the two variables. The regression equation follows the trend of the data.

There are many types of equations that you will study and have studied in algebra. Some examples of these equations are linear, quadratic, cubic, quartic, exponential, and logarithmic. These are also the types of regression equations we can find.

What is the purpose of finding regression equations? We find regression equations so we can visualize the relationship between two quantities in order to make predictions or draw conclusions.

In this activity, you will plot a set of data by hand, plot a set of data on your calculator as a scatterplot, find a regression equation using your graphing calculator, and analyze why you may choose one type of regression equation over another. We will also examine domains, ranges, and functions.

### Getting Started With a Data Set

**The Ideal Gas Law**

The Ideal Gas Law is a thermodynamic equation of state for an ideal gas. An equation of state simply relates the state properties of a substance to the amount of that substance. Specifically, the ideal gas law relates the pressure (\( P \)), volume (\( V \)), and temperature (\( T \)) to the amount (\( n \)) of an ideal gas.

A one liter tank containing one mole of an ideal gas was heated to the temperatures listed in row one. At each temperature, the pressure from a pressure gauge affixed to the tank was recorded in row two.

<table>
<thead>
<tr>
<th>( T ) (temp. Kelvin)</th>
<th>300</th>
<th>310</th>
<th>320</th>
<th>330</th>
<th>340</th>
<th>350</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) (pressure atm)</td>
<td>24.7</td>
<td>25.3</td>
<td>26.4</td>
<td>27</td>
<td>28</td>
<td>28.5</td>
</tr>
</tbody>
</table>

Note: A “mole” is defined as the number of carbon atoms in 12 grams of carbon-12, which is approximately 6.022X10^23.
Plotting the Data by Hand

You will be plotting the data on the grid shown below, graphing temperature on the x-axis (temperature is the domain and the independent variable). You will be graphing pressure on the y-axis (pressure is the range and the dependent variable). Recall that the domain is the set of all possible values of an independent variable of a function and the range is the set of all possible values for a dependent variable of a function.

Step 1: Label the x-axis and the y-axis of your grid. Let’s start with the x-axis. We will graph the temperature on the x-axis.

- What is your smallest x-value (temperature value)?
- What should be your smallest x-value on your x-axis? (Make sure that your smallest value is on the x-axis. You need to have your x-axis go lower than the smallest value.)
- What is your largest x-value (temperature value)?
- What should be your largest x-value on your x-axis? (Make sure that your largest value is on the x-axis. You need to have your x-axis go beyond the largest value.)
- Label these values on your grid below.

- What is your smallest y-value (pressure value)?
- What should be your smallest y-value on your graph? (Make sure that your smallest y-value is on the y-axis. You need to have your y-axis go lower than the smallest value.)
- What is your largest y-value (pressure value)?
- What should be your largest y-value on your y-axis? (Make sure that your largest value is on the y-axis. You need to have your y-axis go beyond the largest value.)
- Label these values on your grid below.

Step 2: Now that your grid is labeled on the x- and y-axis, plot the data points from the table onto the grid.
**Step 3:** Look at your data. If you were to find an equation of a function that best fits this data, would this equation be a line? Why or why not?

Would this equation be a parabola? Why or why not?

Would this equation be an exponential? Why or why not?

**Make sure to remember that you are trying to find the equation that would “best fit” the graph. Not the equation that exactly fits the graph.**

If you have plotted the data correctly, the equation that will best fit this data is a line. Use a straight edge to sketch a line that “best fits” the data. To sketch the line that best fits the data, sometimes it is better to use a clear ruler. When sketching the line of best fit you are trying to draw a line that comes as close to as many points as possible and balances the overestimates and underestimates. In other words, the line should have the same number of points above it as it does below it.

Explain why your line best fits the data.

Once you have drawn this line, you can write the equation by using any two points on it to find the slope. Then use one of the points on the line, the slope you found and the point slope equation of a line to write the equation.

What is the equation of the line that best fits the data?
Finding the Equation for the Line of Best Fit Using Your Calculator

Now, you will enter this data into the calculator and use it to find the equation of your line of best fit. We will compare this with the equation that you found above. Please use the instructions on the document labeled Regression Instructions on the Math 143 MLC webpage.

First, find the set of instructions that says: “Graphing a set of data points on the TI-83” (making a scatterplot). Follow these instructions to plot the points on your calculator.

Once you have plotted the data on your calculator and you see a plot similar to the one that you sketched, observe the scatterplot.

Does your graph on the calculator look like the graph that you sketched? Why or why not? If they do not look the same you might want to explore reasons such as window size, domains and ranges, sketched points correctly or entered data correctly in the calculator.

Once you have resolved any differences between your sketch and the calculator, look at the graph on your calculator. Do you think that a line best fits the data? Why or why not?

Do you think that an exponential equation best fits this data? Why or why not?

Do you think that a quadratic best fits this data? Why or why not?

If the data was entered and plotted correctly, it should most resemble a line just like in your sketched graph. Now you want to find a line that is a good "fit." In the plotted data we showed one method to find this line. Using your calculator we will use another method to find this line.
When you use your calculator to find this line, the line is sometimes called the "least-squares" regression line. Follow the steps given in the regressions document labeled “finding the linear regression equation” to find the equation of this line.

What is your equation of the line of best fit?

In your calculator, you use $x$ and $y$ as your variables. In chemistry, our variables are $P$ for pressure and $T$ for temperature so rewrite your equation using $P$ and $T$ instead of $x$ and $y$. What is your equation using $P$ and $T$?

Finally, you can use a linear regression to create a mathematical model that represents the relationship between two variables. Use your line of best fit to predict what the pressure will be when you raise the temperature to 450 Kelvin. What will be the pressure when the temperature is lowered to 250 Kelvin?

In the Ideal Gas Law equation is $P = \frac{nRT}{V}$

where

$P =$ pressure of an enclosed group of gas molecules in atmospheres (atm)

$V =$ volume of an enclosed group of gas molecules in units of liters (L)

$n =$ moles of gas (1 mole of gas = 6.022X10^23 molecules)

$T =$ Temperature of gas molecules in units of Kelvin (K)

$R =$ Gas law constant = 0.08206 (L*atm)/(mol*Kelvin).

In our situation we left $n$ and $V$ constant. We left both of them as 1. Thus the Ideal Gas Law equation when $V$ and $n$ are 1 is $P = RT$. When we wrote our equation above, we left $P$ and $T$ in our equation as our $x$ and $y$. Thus your above equation should look very close to $P = 0.08206T$.

**Exercises**

For the data sets below:

- determine the domain and range
- plot the data on the graph making sure to label your “window” on your paper
- sketch a line of best fit
- enter the data into your calculator
- find the line of best fit
- compare your sketched line of fit to that of your calculators found line of best fit
1. A student is hanging masses from a spring and measuring the resulting stretch in the spring. The following table shows the students' collected data.

<table>
<thead>
<tr>
<th>m (mass in grams)</th>
<th>15</th>
<th>25</th>
<th>35</th>
<th>45</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>x (stretch in cm)</td>
<td>7.2</td>
<td>10.6</td>
<td>14.3</td>
<td>21.6</td>
<td>24.6</td>
</tr>
</tbody>
</table>

2. A student is hanging masses from a spring and measuring the resulting stretch in the spring. The following table shows the students' collected data.

<table>
<thead>
<tr>
<th>m (mass in grams)</th>
<th>2</th>
<th>6</th>
<th>9</th>
<th>15</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>x (stretch in cm)</td>
<td>15.3</td>
<td>21.0</td>
<td>25.1</td>
<td>33.4</td>
<td>43.2</td>
</tr>
</tbody>
</table>