

Goals

The goals of this experiment include gaining an understanding of data analysis in the physical sciences and the importance of graphical analysis.

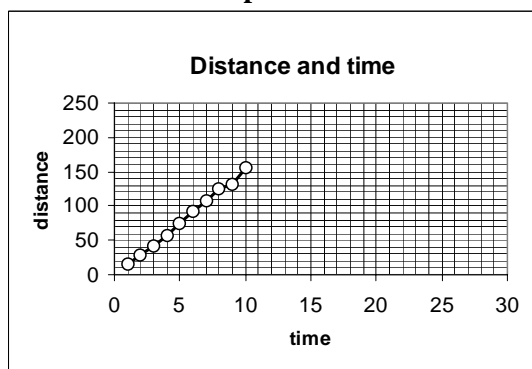
Equipment and Materials

Transparent metric ruler, pencils, PC with graphical analysis software, graph paper, USB flash drive.

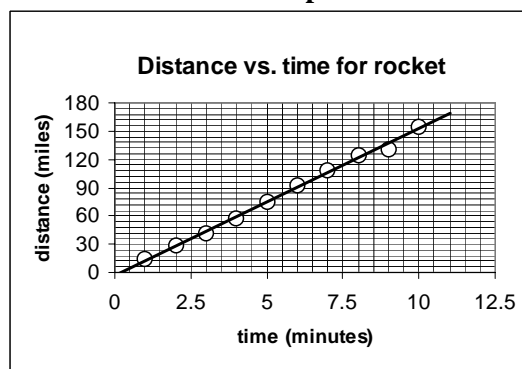
Discussion

Graphs are often used to represent numerical observations (measurements). For example, consider the motion of a rocket. We could measure the distance that the rocket has traveled over time. If we made many measurements it would be difficult to analyze the meaning of the data. In order to understand the data we could produce a graph that shows how the distance traveled depends on time. Consider the two graphs shown below and note their similarities and differences:

Graph 1



Graph 2



Graph 1:

1. The title doesn't give much information about what the graph is showing.
2. The axes are labeled, but without units. Does the time shown mean seconds or minutes or years?
3. Only a small portion of the graph paper is used. This bunches everything in one corner.
4. A straight line wasn't used as the line of best fit. Connecting the dots doesn't indicate whether there is a simple mathematical relationship between distance and time.

Graph 2:

1. The title explains what is shown in the graph.
2. The axes are labeled with units drawn in parentheses.
3. Round numbers are used to give an idea of scale on both axes. The numbers follow a linear pattern and are regularly spaced.
4. Most of the graph area is used.
5. The data points are clearly shown with open circles.
6. The line of best fit is shown.

Overall, Graph 2 is superior to Graph 1 because of the differences listed.

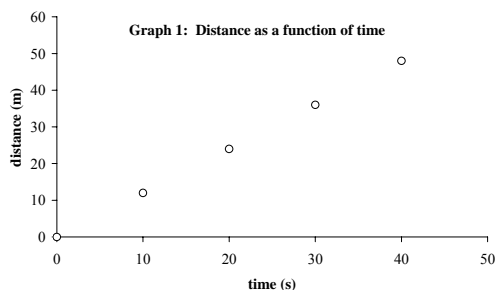
Linear Regression

It is often necessary to make predictions or draw conclusions based on experimental data. Graphical analysis can often help in this goal. Consider the data shown in Table 1:

Based on the pattern of this data we can estimate that at 50 seconds the distance should be 60 meters because the distance appears to increase by 12 meters every 10 seconds. However, this type of estimate is fairly difficult to make with most sets of data.

A graphical representation of the data can sometimes make the process of estimation easier. To predict the distance at 50 s we could draw a graph of the data

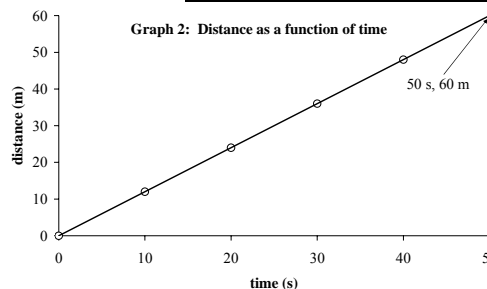
shown in the table. In this graph we plot distance on the vertical axis and time on the horizontal axis (Graph 1).



Notice that

each point on the graph represents a paired set of data (one for distance and one for time). In the graphs that we will draw, it is always the case that each point on the graph corresponds to a paired set of data (for example, pressure and temperature).

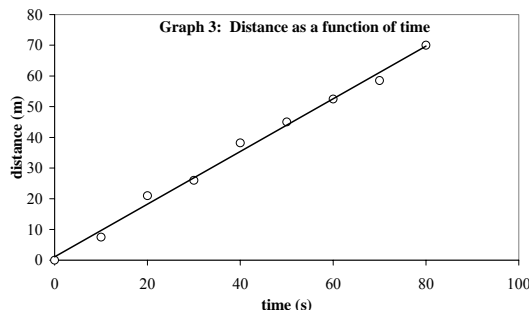
Time (seconds)	Distance (meters)
0	0
10	12
20	24
30	36
40	48
50	?



One variable often depends on another. For example, the distance traveled depends on how long you've traveled. In this case, distance depends on time. In general, when one variable depends on the other, the one that is independent is drawn on the horizontal axis. The one that is dependent is drawn on the vertical axis. For Graph 1 we would say that it is a graph of "distance versus time" because distance is the dependent variable and time is the independent variable.

To project the value of distance when the time is 50 seconds is made easier by drawing a straight line through all the points and extending it to 50 s. When we do this we get Graph 2. By looking at the line, it's easy to see that at 50 seconds, the distance is 60 meters.

The process of using a line to project the values of distance for given times is often called linear fitting. The line is drawn based on the actual data values. But where there are no data values we can use the line to determine what they would be. In the data for Graphs 1 and 2, the line fits the data perfectly, meaning that the line overlaps with all the data points. This rarely happens with real experimental data. Usually, there is no line that perfectly fits over all the data points. So we must draw a line that comes as close to all the points as possible. This line is often referred to as the line of best fit.



The line of best fit can be drawn with a pencil and a ruler. Graph 3 shows the case in which there is no line that can perfectly fit all the data points. So we draw a "line of best fit." Even though the line of best fit doesn't perfectly fit the data, we can still use it to project data values we haven't measured. There is a

mathematical analysis that can be used to calculate the line of best fit. It's not based on estimation with a pencil and a ruler. It's based on calculation and is called linear regression. The calculation is extensive and is usually done using computer software.

Equation for a Line

The equation for a line is given by $y = mx + b$ where y is the data that corresponds to the vertical axis, x is the data that corresponds to the horizontal axis, m is the slope of the line, and b is the y-intercept (the value of y when x is equal to zero).

Sample Calculation

Suppose that a graph of distance versus time (y versus x) gives a slope equal to 10 m/s and a y-intercept equal to 5 m. If the distance is equal to 55 m what is the time equal to?

To solve this we use the equation for a line:

$y = mx + b$ where y is distance and x is time.

distance = m (time) + b

We know the slope, y-intercept, and distance. So we can solve for the time:

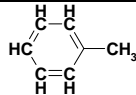
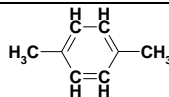
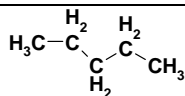
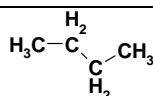

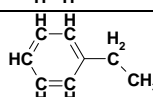

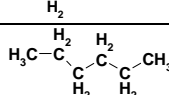
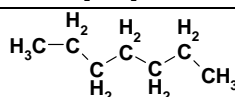
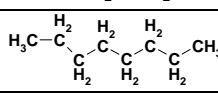
$$\text{time} = \frac{(\text{distance} - b)}{m} = \frac{(55\text{ m} - 5\text{ m})}{10 \frac{\text{m}}{\text{s}}} = \frac{50\text{ m}}{10 \frac{\text{m}}{\text{s}}} = 5\text{ s}$$

Notice that the units seem reasonable because $\frac{\text{m}}{\left(\frac{\text{m}}{\text{s}}\right)} = \text{m} \left(\frac{\text{s}}{\text{m}}\right) = \text{s}$

Source of Data

A report from 1971* showed that soil can act to destroy carbon monoxide (CO). Carbon monoxide is an air pollutant produced by automobiles and is produced from the incomplete combustion of gasoline in cars and trucks. Combustion is a type of chemical reaction in which oxygen reacts with a hydrocarbon, often forming water and carbon dioxide. Gasoline is a mixture of hundreds of compounds called hydrocarbons (compounds that contain carbon and hydrogen). Most of these compounds contain between five and twelve carbon atoms. Some common hydrocarbons found in gasoline are shown in Table 2.

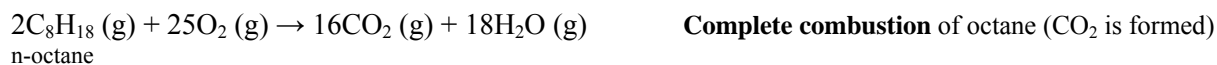
Table 2: Common hydrocarbons in gasoline

Name	Amount in Gasoline(%) ¹	Chemical Formula	Chemical Structure
toluene	0-35	C ₇ H ₈	
xylene	0-25	C ₈ H ₁₀	
n-pentane	0-20	C ₅ H ₁₂	
n-butane	0-12	C ₄ H ₁₀	
benzene	0-5	C ₆ H ₆	
ethylbenzene	0-5	C ₈ H ₁₀	
cyclohexane	0-5	C ₆ H ₁₂	
n-hexane	0-5	C ₆ H ₁₄	
n-heptane	0-2	C ₇ H ₁₆	
n-octane	0-1	C ₈ H ₁₈	

The amounts given represent a range in percentages. Gasoline contains a variable amount of substances based on the processing and the source of the material.

* Soil: A Natural Sink for Carbon Monoxide, Inman, Ingersoll, and Levy, Science, vol. 172, 1971, p. 1229.

One component of gasoline is n-octane, C₈H₁₈. The chemical equations below represent the incomplete and complete combustion of octane:



When gasoline, oil, or diesel fuels are burned, incomplete combustion can result in concentrations of carbon monoxide as high as 50-100 parts per million (ppm) in the air.* This level of CO is harmful to some people. Furthermore, the amount of CO released by cars and power plants is, in theory, enough to cause its concentration in the environment to increase rapidly. But the actual average concentration has not increased dramatically (0.07 ppm). An explanation for this discrepancy is that some species of bacteria, located in soil, can destroy CO, perhaps by oxidizing CO to CO₂:



The action of soil bacteria on carbon monoxide can be tested. In one experiment, a sample of soil was placed in a container with a sample of air containing 100 micrograms of CO per gram of air. Over the period of a few hours the CO content in the container was measured. Table 3 shows some of the data from these measurements.

Table 3: Carbon Monoxide Concentration as a Function of Time

**The units discussed in the text (parts per million) are concentration units. The unit of parts per million (ppm) is the ratio of the amount of the substance of interest (carbon monoxide) compared to some other substance (in this case, air). In a 100 μg sample of CO and 1 g of air the ratio is:

$$\frac{1\mu\text{g CO}}{1\text{g air}} = \frac{10^{-6}\text{g}}{1\text{g}} = 1 \text{ part per million}$$

So 100 μg CO per gram of air is 100 ppm.

*1 μg = 1 microgram = 1 x 10⁻⁶ grams.

Time (minutes)	CO content (μg)	Time (minutes)	CO (μg)**
0	1320	74.0	777
8.5	1225	91.5	609
17.5	1183	108	477
34.5	1023	127	368
53.0	887		

In this workshop you will plot the data in the table above and draw the line of best fit to determine how long it takes for the amount of carbon monoxide to reach zero concentration. The assumption we will make while performing this data analysis is that carbon monoxide concentration decreases in a linear fashion with time. When we generate the graph we will see that this is the case.

Procedure

Students work individually on this experiment. This workshop will take place in a computer laboratory. Bring a USB flash drive for saving data and a ruler.

In this experiment, the data has already been generated and is shown in Table 3. We will generate two graphs in this workshop. The first graph (Part 1) will be drawn by hand using graph paper. The second graph (Part 2) will be generated on a computer using graphical analysis software.

Part 1: The hand-drawn graph

Several sheets of graph paper are included at the end of this experimental procedure. Use these sheets for the hand-drawn graph.

1. General rules for drawing graphs. Take a look at the graphs shown on page 9.
 - Draw each data point as a small circle (\circ). If more than one kind of data is to be drawn on the same graph, you can also use a triangle, square, or other figures.
 - Use as much of the graph paper as is possible.
 - Always label the axes of the graph.
 - Title the graph with information that indicates what the graph represents.
 - Numbers should be equally spaced.
 - Use round numbers on the axes.
2. Choose scales for the graph that allow a time range of 0-200 minutes and a CO content range of 0-1400 μg . Label the axes being sure to include units. Time should appear on the horizontal axis and carbon monoxide content on the vertical axis. In graphing, the horizontal axis is often referred to as the x-axis, and the vertical axis as the y-axis.
3. Using the data from Table 2, graph the points for CO content "versus" time. You should have nine points on the graph.
4. Using a ruler, draw the line of best fit through the points. The way you do this is somewhat arbitrary. But you should position your ruler so that the line you draw has about as many points on one side of the line as on the other side. A more sophisticated and accurate method for finding the line of best fit, called linear regression, involves calculations that are best performed by a computer (Part 2 of the experiment).
5. Extend the line of best fit (as a dotted line) until it crosses the x-axis, and estimate from the graph how many minutes (x) it would take for all of the CO to be removed. The extension of the line beyond the range of the data is called extrapolation.
6. Once you have completed the hand-drawn graph, show it to your instructor. Once you are satisfied with the graph move on to Part 2.

Part 2: The Computer-generated Graph

1. Open the application Graphical Analysis.

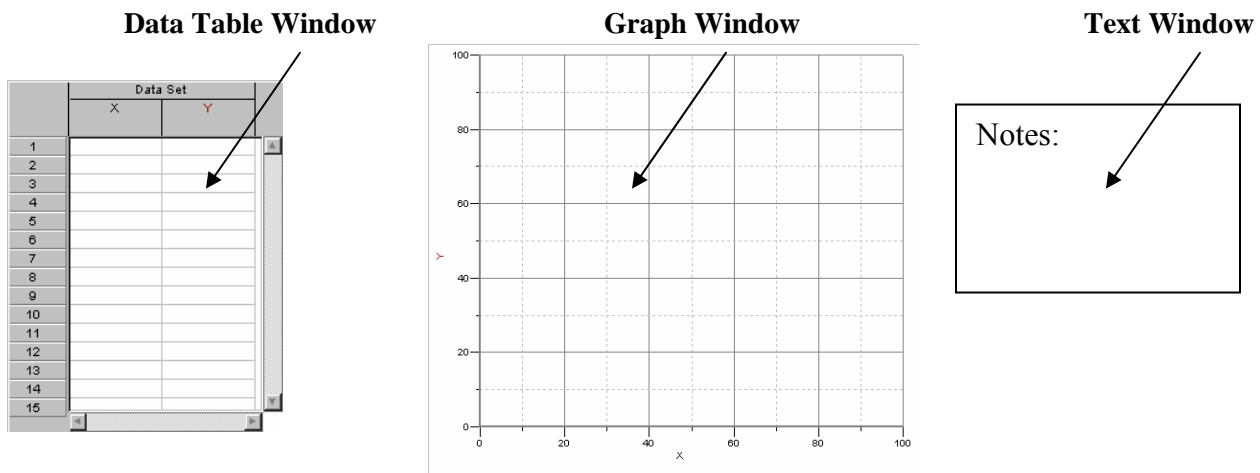
Note: The version of software used in this laboratory is subject to change. If a different version of the software is being used than is described in this procedure you may need to explore the menu options of the program to access certain functions. Be sure to save the document to a flash drive frequently. This will reduce any losses if the computer you are using happens to crash.

Graphical analysis can be performed by hand with graph paper, a pencil, and a ruler. This is a quick way to analyze data that follows a linear pattern. However, the method for finding the line of best fit using a ruler is, at best, an approximation. Alternatively, the line of best fit can be calculated mathematically to great accuracy. The process used to accurately determine the best line of fit through calculation is called **linear regression**. However, performing these calculations by hand are time-consuming and beyond the scope of this course. Computers can be used to quickly and easily determine the line of best fit using linear regression. Part 2 of this workshop will provide experience in using graphical analysis software to find the line of best fit for the data graphed by hand in Part 1.

Using the Graphical Analysis software, graph the data as you did in Part 1. Follow the instructions provided below to complete this assignment.

There are 3 windows that you will use in this program.

1. **Data Table Window:** This window allows you to enter data so it can be graphed.
2. **Graph Window:** This window will show the data in graphical form.
3. **Text Window:** This window allows you write notes concerning the data and graph.



1. Giving a Title to the Graph

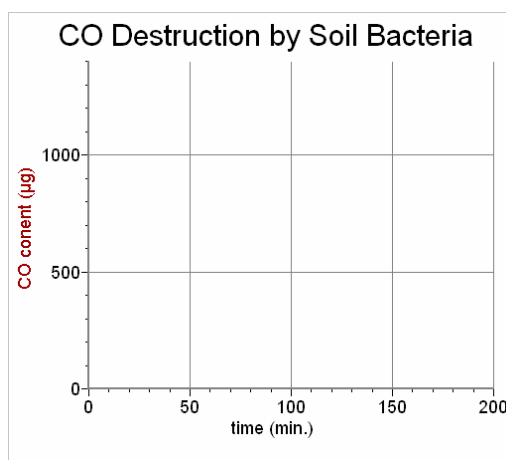
- Highlight the graph window with the mouse.
- Double-click somewhere on the graph window and a window will appear entitled **Graph options**.
- Enter the title for the graph: "CO Destruction by Soil Bacteria" and click **Done**.

2. Labeling the Axes

- Highlight the data table window with the mouse.
- Move the cursor over the X in the data set area.
- Double-click on the X and a window called **Column Options** will appear.

- Give a name to the axis: *time*
- Using the mouse, highlight the units box and enter *min.* as the units and click **Done**.
- Delete the contents of the short name box.
- Do the same for the y-axis using *CO content* as the name.
- The units for the y-axis are micrograms (μg). To enter the symbol μ click once on the units box and then select μ by clicking on **insert symbol** and selecting the μ symbol. Type g to finish the units.

Note that the graph window now shows the new axes labels and units that were just entered.



3. Changing the Scale of the the Graph Axes

- Select the graph window and then double-click somewhere on the graph.
- Select the tab **Axes Options**.
- Select **Scaling** for the Y-axis and change the setting to **manual**.
- Change the limits of the Y-axis to 0-1400.
- Similarly, change the limits for the X-axis to 0-200 min.
- Select **Done** and the graph window should now look something like that shown in the image shown above.

4. Entering the Data

- Data is entered in the data table window. As data is entered points will appear in the graph window.
- Select the data table window.
- Move the cursor to the first data cell under the time column.
- Select the cell and enter the data point into the cell (**0.0**) and select Enter.
- The cursor will move to the first CO Content cell. Enter 1320 and press Enter.
- Continue entering all the data through 127 minutes. If you enter an incorrect number then select the cell and re-enter the correct number.

Table 1: Carbon Monoxide Concentration Over Time

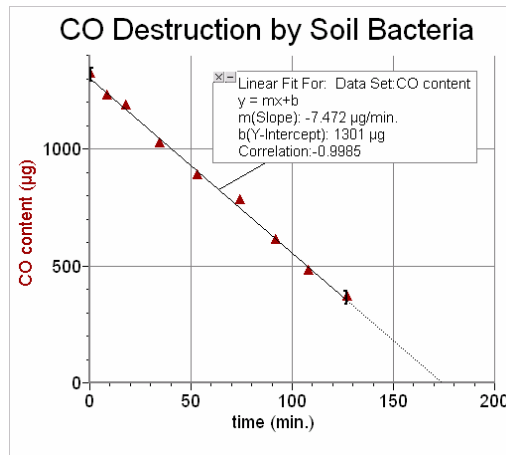
Time (min.)	CO content (μg)	Time (min.)	CO content (μg)
0.0	1320	74	777
8.5	1225	91.5	609
17.5	1183	108	477
34.5	1023	127	368
53.0	887		

5. Data Analysis

The data should now be represented on the graph. The concentration of carbon monoxide decreases over time as a result of the bacteria in the soil. Analysis of the graph involves linear regression, the line of best fit.

- To generate the line of best fit we'll use the graph window and the **Analyze** menu.
- Select the graph window.
- Double-click anywhere on the graph and the **Graph Options** window will appear.
- Select the tab **Graph Options**. Click the small box to the left of **Connect points** to de-select this option. Select **Done** and the connecting lines should no longer appear on the graph window.

- Select the menu **Analyze**.
- Scroll down to and then select **Curve fit**.
- **Curve Fit** allows you to choose one of many different fits to the data.
- Select **mx + b** (linear fit).
- A new menu at the bottom of the screen will prompt you to decide whether the fit is OK or not. If you think the linear fit is adequate then left click once on **OK-Keep fit**.



to a

The small window in the middle of the graph can be moved out of the way by dragging it to the side. The window contains the general equation to a linear fit with the slope (m) value and y-intercept (b) values indicated. The **Correlation** is a measure of how close linear fit the data is. Values near 1 indicate that the data follows a linear relation.

6. Determining the time at which the CO content is zero µg.

The equation for the line should be shown in the dialog box on the graph window. The equation is: $y = mx + b$ where y is the CO content and x is the time. We want to find the time when CO content is zero. To determine this we can re-arrange the equation to solve for time (x):

$$y = mx + b$$

$$mx = y - b$$

$$x = \frac{y - b}{m}$$

Since x is time in this case we can use the slope and y-intercept values from the graph and set y equal to zero. This gives:

$$x = \frac{y - b}{m} = \frac{0 - b}{m} = \frac{-b}{m}$$

7. Entering Text in the Text Window

- Select the text window and type the following information in the window:

Text Window

First Name	Last Name
Chem. 11	Section _____
Time for zero CO content _____ min.	

- The time for zero CO content is found by using the linear equation and the values found for the slope (**m**) and y-intercept (**b**).
- Once you have shown your graph to your instructor, print the entire screen and turn it in with your report. Be sure to save the file to your flash drive.

Questions

Answer these questions and turn them in with your report.

1. Using the data below construct two graphs using the Graphical Analysis software. In the first graph plot volume (the units are in cubic inches) versus pressure (the units are inches of mercury). In the second graph plot volume versus $\frac{1}{\text{Pressure}}$. You'll need to calculate $\frac{1}{\text{Pressure}}$ for each value of pressure. Compare the two graphs. Does it appear that either set of data could be fit with a line? If so, construct a line of best fit.

Volume (in ³)	Pressure (in Hg)	$\frac{1}{\text{Pressure}}$ (in Hg) ⁻¹
118	6.0	
87	8.0	
71	10.0	
59	12.0	
44	16.0	
35	20.0	
29	24.0	

2. Consider the volume and temperature data for a gas given below. Graph the volume versus temperature. At what temperature is the volume equal to zero liters?

Temp (°C)	Volume (liters)
-150	2.261
-100	3.1797
-50	4.0987
0	5.0177
50	5.9367
100	6.8557

3. Construct a graph of temperature in degrees Fahrenheit versus temperature in degrees Celsius using a computer and the following data:

- The temperature that water freezes
- The temperature that water boils
- The temperature of the human body
- Any two other interesting temperature values

4. Submit a one-paged (typed) report that discusses the following topics:

- a. What were the goals or objectives of the experiment
- b. What were your results as they relate to the goals and objectives.
- c. Discuss the quality of your results.

Observations and Notes
Experiment 2: Workshop in Graphing Experimental Data

Date _____



