Case Study
Effect of Neurotoxin Pesticides on Lung Vital Capacity*

The average lung vital capacity as measured with a spirometer is about 4.2 liters for a healthy average-size adult male. Farm workers were accidentally exposed to a cholinesterase inhibitor pesticide being sprayed on strawberry fields. The workers exhibited respiratory distress. To assess the damage to their lungs, they were given vital capacity tests. The pesticide exposures and vital capacities for 16 workers are shown below.

The potential effect of the neurotoxin pesticide on vital capacity has been established. Pesticide exposures that result in a vital capacity of 2 liters or more have been determined to be safe.

* Data have been fabricated for this example case study. You might want to do an internet search to read about cholinesterase inhibitors and their potential effects.

Effect of Neurotoxin Pesticides on Vital capacity

Farm Worker	Pesticide Concentration (ppm)	Vital Capacity (liters)	Farm Worker	Pesticide Concentration (ppm)	Vital Capacity (liters)
1	150	3.9	9	1495	2.8
2	225	3.8	10	1050	3.1
3	375	4.3	11	0	4.4
4	950	3.7	12	705	4.0
5	1005	1.9	13	945	3.1
6	670	4.1	14	1380	1.7
7	245	4.6	15	845	3.6
8	1155	2.2	16	450	2.9

Discussion

1. Which, if any, of the farm workers were exposed to a toxic dose of the cholinesterase inhibitor pesticide?

2. Using the data from the table above, graph the pesticide concentration vs. vital capacity for the 16 farm workers on the template provided. Choose the type of graph (bar graph, line graph, scatterplot) you will use.
    *   Explain why you chose that type of graph.

    *   What is the independent variable? _____.
    *   What is the dependent variable? _____.
            Be sure to include units for each variable.

    *   Describe how you decided what scale to use on each axis.

3. Draw on your graph a best-fitting regression line for your data. You may just "eye-ball" the regression line. Be sure you have read the information on regression lines at the end of this exercise before you complete questions 3 – 5
.
4. Estimate the slope of the best fit line you drew. _____.
    *   Which points on the line are you using to estimate the slope? _____.
    *   Is your slope positive or negative? _____.
    *   What are the units of the slope? _____.

5. Estimate the y-intercept of your best fit line. _____.
    *   What concentration of pesticide exposure does it correspond to? _____.
    *   How does it relate to the average vital capacity of an adult average-sized male?

    *   Provide an explanation for any difference you find.

More on Regression Lines
When we study metabolic processes such as respiratory rate, heart rate, photosynthesis, cellular respiration or osmotic activity, we need to measure the variable being tested (e.g., oxygen consumption, lung capacity, carbon dioxide production) over time. To best evaluate the data obtained we should graph our results. We obtain a scatterplot when the discrete data collected are plotted as points on a graph. These points rarely fall on a straight line. But sometimes, they are close to a straight line indicating a linear relationship between the varables.

To find the best fitting straight line for the data in the scatterplot, we use linear regression. There are formulas for calculating the "y = mx + b" form for a regression line, but calculators and computers (such as the data recorders we will use in lab) find this information for us much more quickly than doing it "by hand." Thus, we need to know how to interpret the information the computer (or data recorder) provides us.

When the data recorder graphs a regression line, it also displays a box that reminds you that the equation of a line is y = mx + b. The box also contains the m and b values. The m value is, of course, the slope of the regression line, and in most of our labs it's the rate of change we are interested in.

For example, if the data recorder displays "m: 0.15   b: – 0.005", this means that the equation of the regression line is y = 0.15 x – 0.005. Since the slope is 0.15 and we know this is a rate of change, we can say that the y variable will increase by 0.15 units for a one-unit increase in the x variable. The units here depend on the labels on the x and y axes.

Sometimes we have a scatterplot of data points, and no data recorder to calculate the regression line for us. One quick alternative is to "eye-ball" a best fitting regression line directly on the scatterplot. We can then estimate the slope m and y-intercept b by looking at the line drawn. Recall the y-intercept is the value where the line crosses the y axis. To find slope, pick two points on the regression line and then find "rise over run".

Some things to keep in mind:
- A regression **line** is a good fit for the data only if the points represent a **linear** relationship.
- The regression line will not pass through all points. In fact, it may not pass through any of the data points! It should, however, be a close fit to most of the points in the scatterplot.
- If the points do not exemplify a linear relationship, there are other regressions (quadratic, exponential) that may fit the data better. (They are not a part of this class.)