Statistical Analysis
Microsoft's Excel program allows you to easily do the simple statistical analysis you need to examine the data you have collected.

Getting started
You can skip this section if you are already familiar with spreadsheets.
1. Open the Excel program from the toolbar or under programs in the start menu. The Excel screen will appear; numbered rows down the left, lettered columns across the top. Each page is considered a worksheet.
2. Each one of the little squares is called a cell. Each cell is identified by a column and row coordinate. The little drop box in the left corner just above the worksheet indicates which active cell you are in. A1 will appear there when you open a new spreadsheet. Click on any other cell to observe how it tracks your position on the work sheet. You can also type a cell identification and the cursor will move to that cell.

Entering data for analysis
You have collected a lot of data. We initially recorded this data manually. You need to enter it into the spreadsheet so the computer can perform the analysis. The worksheet can be structured in a variety of ways. It is often simplest to manually enter the two sets of data you are interested in, rather than have Excel sort the data for you.
3. It is good housekeeping to save the far left column and the top row for identifying your data sets. Excel recognizes the difference between numbers and words.
4. Move you cursor to B1 and type in an identifying heading for your first set of data. You can use the arrow key to move down the column.
5. Enter all the data for your first data set down that column. Be sure to move into an empty square after your last data entry so that value will be recorded.
6. Next move your cursor to C1 and enter your second set of data down that column.
7. Do not leave any blank spaces in your columns. The program sees them as zeros.

Be sure all the data you need is entered and correct. It must be displayed in the correct significant figures. If you measured something to the nearest whole mm, you must report and analyze your data to the nearest whole mm. The computer can calculate numbers out to the bazillionth place. This would imply a precision to your measurement that you do not have. You can fix this on the toolbar. There is a shortcut button that increases and decreases the number of decimal points displayed. You can also go to Format on the toolbar- select Cells- select Number sheet- under category select Scientific and a dropdown box will appear. Set the number of decimal places appropriate for your data.

Statistical Analysis
8. Double click on the ***f*x** or function button either on the toolbar or in the formula box. A drop down menu will appear.
9. Scroll through the Function category box on the left until you see Statistical. Highlight Statistical and click on that. The Function name box on the right will change.

There are an incredible number of tests to choose from. The trick is choosing the correct test for the type of data you have collected.

Mean or average

We have a series of numbers that we collected. Now we need to make some sense of them. There is a low and a high value or range of numbers that we collected. This does not tell us much about the numbers in between. We need to calculate the mean or average of the values. This is the sum of all the values divided by the number of data points. Fortunately the computer does this very easily.

10. You will need to first indicate a cell in your spreadsheet where you want this information to appear. One way to do this is to pick a cell below your last number in the data column. It's a good idea to leave an empty cell above where you want your results to appear. Use your blank first column to type in some identifying word.
11. Select AVERAGE from the Function name box. Highlight and click.
12. A new drop down box appears, usually right where your data is. Move the arrow to a gray area on the box and drag it to where you want it.
13. Number 1 and Number 2 appear in the box. You can enter the cell identifier for your first number in the Number 1 box then enter a colon and type in the cell name for the last number in your column. Alternately you can move your cursor to the first number, hold down the button and drag the cursor down to the last value. The computer will automatically enter the correct cell numbers.
14. Hit the OK button and the calculated value will be recorded in whatever cell you left the cursor in before you started the analysis. It will also appear in the bottom left hand corner of the drop box screen.
15. Repeat this procedure for your second set of data. Be sure your answers are displayed to the appropriate level of accuracy.

Standard Deviation

The mean gives you one piece of information about your data. It does not tell you how much your values vary from the mean. 4, 5, 6 have a mean of 5. So does 1, 5, and 9! The closer all your data is to the mean the better the mean is at representing the information you collected and the more accurate the conclusions you draw from your data. Standard deviation is the measurement of variance of your data from the mean, both above and below it. It is indicated by mean $\pm$ standard deviation. The smaller the standard deviation the "tighter" your data, the more accurate the mean is at representing your data.

16. Pick another cell where you want the calculation of the Standard Deviation to appear. Be sure you are in an empty cell.
17. Return to the Statistical menu, scroll down till you see STDEV, highlight and click. Look for a short description of the calculation the computer is doing in the gray drop-box.
18. Either manually enter the cells in your data set or highlight your data. Hit OK.
19. The standard deviation of your data will be entered into the empty cell.
20. Repeat the process for your second set of data.

The smaller your standard deviation is, the more representative the calculated mean of your data.

Student t-test

You now have some useful information about your data. If the means of your two samples are very far apart you can be pretty confident in saying the two data sets are different. But because biological systems are so variable we most often are looking at data that have small differences. Statistics is a mathematical tool that allows you to say with a certain level of predictability and confidence if the difference you observe in your data is due to chance or reflects a real, statistically significant difference in your data.

We collected two sets of data in this experiment. The type of data we collected is considered parametric. We could have measured it infinitesimally small and infinitesimally large. Non-parametric data is yes/ no, ranked or categorical type data. We also assumed that our data follows a normal curve, or Gaussian distribution. Some of it was small, some large but most of it was close to the mean. The t-test is an appropriate and simple test for analyzing the means of two parametric data sets.

21. Be sure to leave your cursor in the empty cell where you want your data to appear.
22. Return to the ***f*x** key; select Statistical and scroll down to TTEST. Highlight and click.
23. A new drop box will appear. Array 1 refers to your first data set. Highlight and enter your data as before.
24. Put your second set of data in the Array 2 box.
25. Next you need to tell the computer if it is a 1 tailed or 2-tailed distribution. Enter 2 because our data is evenly distributed above and below a mean. Think of how a graph of the normal curve looks.
26. We also need to tell the computer what type of information we collected. Paired refers to information collected on the same subject such as before and after intervention and must be entered appropriately in the respective columns. Most of the data we collected is un-paired; there was no matching of the two groups. It also had equal variance. Enter 2 in this box.
27. The computer calculates and displays the p-value. What this number says is the probability that the difference in the means of your two samples is not due to chance. 95% is the generally accepted level of confidence. If the p-value calculated by the t-test is less than 0.05, then at the 95% level of predictability the means of your data is statistically significantly different. If your p-value is greater than 0.05, then your data means are not statistically significantly different and the differences you see are most likely due to chance.