Comparing the Means of Independent Samples
The t-test

The t-Test is a test for differences between means of independent samples. It allows us to use statistics to say whether two means are different or the same, rather than simply comparing the data and "guesstimating" if they are the same. There are several things that must occur for this test to work.

- The sample size must be large (above 25) and normally distributed. A normal distribution is one that fits well to a bell-shaped curve, with the same number of samples on either side of the mean. Otherwise a different (non-parametric) test is used.
- The distribution of observations must be continuous rather than discrete. This means that there are theoretically an infinite number of measures that could be taken. For instance, the height of all students in the class is continuous, but determining how many males versus females there are is discrete (there are only two possibilities).
- The samples must be independent of each other, which means the data must be collected from different study populations (two different classes) rather than the same one sampled multiple times (the same class tested on two different days).
- The variances must be equal. The variance is the statistical measurement of the dispersion of the data around the mean. Variances can be compared using the F-test. We will not do this today, but instead assume equal variances.

For the t-test, you have a null hypothesis (H_0), which is typically that the means you are comparing do not differ. This is stated $H_0 : X_1 = X_2$, that is the mean of sample 1 equals the mean of sample 2. This hypothesis is assumed to be true until you can prove otherwise.

Your alternate hypothesis (Ha) is that the means do vary, which is stated $Ha : X1 \neq X2$. This is called a 2-tailed test, since you have no prediction about which mean is the larger. If you do, it is a 1-tailed test and the p-value is calculated slightly differently. When calculating the t-statistic, you calculate a t and the degrees of freedom (d.f.), then use a chart to find the p-value. The p-value is essentially the chance that you have incorrectly rejected the null hypothesis (and thus accept the alternate hypothesis). A p-value of 0.05 is the largest p value that is considered significant (less than 5% chance that you incorrectly rejected the null hypothesis). The smaller the p value the better your chance of not incorrectly rejecting the null hypothesis.

Terminology:
S^2 - variance: the amount of dispersion around the mean
$\sum$ - the mathematical symbol for sum
x - any one of your data points
$\sum x^2$ - square each data point, the sum these squares (add up all of the sums)
$(\sum x)^2$ - the square of the sum of all your data points (add up the points, the square the sum)
n - the number of data points you have collected
X - the mean (average) of your data. This is calculated as $\sum x / n$.

Any of the above numbers can include a subscript to indicate which data set from which they should come

X - absolute value - make the calculation and use the value only (not the sign). In effect, this does not change positive results and makes negative results positive

Calculating variance:

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$$

Calculating the t-statistic:

$$t = \frac{X_1 - X_2}{\frac{(n_1 - 1) S^2_1 + (n_2 - 1) S^2_2}{n_1 + n_2 - 2} \quad \frac{1}{n_1} + \frac{1}{n_2}}$$

degrees of freedom (d.f.) = $n_1 + n_2 - 2$

Using this t-value and the degrees of freedom, consult the chart and calculate whether or not you can reject the null hypothesis and accept the alternate hypothesis. To do this, find your degrees of freedom and compare your t-value with the critical t-value. If your t-value is equal to or greater than the critical value, you can reject the null and accept the alternate hypothesis. If your t-value is less than the critical value, then you cannot reject the null hypothesis.