The t-test:

This is a really brief introduction to the t-test. It lays out some of the basic concepts and terminology you will need to know when performing this type of statistical analysis!

- t-tests are for comparing **two groups**. You can use a bunch of t-tests to look at more than two groups, but **for each given t-test**, only two groups are considered.
- as you can see from the formula for the “**calculated t-value**”, the numerator compares the two means, finding the magnitude of the difference between them (in some formulas, the absolute value is taken, so you don’t get a negative number). The denominator has a scary looking formula with a square root and n’s and s’s, but really it isn’t bad at all. This part of the formula gives a **measure of the spread** within each group (variance/standard deviation/standard error).
- calculating a test statistic is really asking yourself the question: is the difference between my group means bigger than the random variation that exists between them?

- **null (H₀) and alternative (Hₐ)** hypotheses are short statements you always state prior to carrying out a statistical test, different than the typical research hypothesis that you are probably used to. They are just one simple sentence each. H₀ assumes that **there is NO significant difference** between your two groups, while Hₐ assumes that **there IS a significant difference**.
- H₀ will always just be something like “**mean of group 1 = mean of group 2**”
- Hₐ will change depending on the context. In a **one-tailed test**, you basically have an idea already in your mind of what you will see, for example, that a new medication will **significantly IMPROVE a certain medical condition**. In the case of a one-tailed test, you will use either > or < in your Hₐ, stating “**mean of group 1 >/< mean of group 2**.” For **two-tailed tests**, you might not have an idea in mind beforehand and just want to see if **any** differences exist between the two groups. For these tests, you will **always be using ≠, never > or <**. For example, “**mean 1 ≠ mean 2**.”

- you also use a big, frightening table to get something known as your **“critical t-value.”** Again, it actually isn’t scary at all! Just a compact way to provide you with lots of information!! This number is kind of like a “cut-off.” For any calculated t-value **greater** than this, we reject H₀ and consider our groups **significantly different**; any calculated t-value **less** than this, and we **fail to reject H₀**, claiming that not enough evidence exists to suggest a significant difference between the two.
  - from this conclusion, you can move onto determining whether your **p-value** is > 0.05, or < 0.05. I like to think of p-values as the “probability that these results are just due to random chance.” Random chance exists everywhere, and there is always the possibility that the difference you observed between your groups is just luck—after all, you will almost never get two means that are exactly the same! There will always be at least some sort of difference, and the p-value lets us know how likely it is that we would observe a difference this large simply by chance.
  - t-calc > t-crit  =>  **REJECT H₀ => p-value < 0.05** (difference between means so big that the probability that this is simply random chance is small, less than 5%)
  - t-calc < t-crit  =>  **FAIL TO REJECT H₀ => p-value > 0.05** (difference between means not enough to overcome the random variation within the groups; chance that seeing this difference is just luck is large, more than 5%)

*Note: Why 0.05? Well, it really just comes down to the fact that a long time ago, people decided that a 5% error rate was what we would settle with. Sometimes you will see numbers other than 0.05 used as the cut-off, but 0.05 is typically the standard cut-off probability. Pretty arbitrary, I know, but that’s just the way things are in statistics sometimes! We need to draw the line for significance somewhere, and 0.05 just happens to be the place we have decided to do that.*
Steps to perform a t-test! 😊

(1) Set up your null and alternative hypotheses (H₀ and H₁). Think about whether it you are doing a one-tailed test (mean 1 >/mean 2) or a two-tailed test (mean 1 ≠ mean 2)!

(2) Collect data and calculate descriptive statistics (mean, variance, standard deviation, standard error, etc...) *Note: I suggest using a computer software such as Excel for some of these, to avoid making arithmetic errors. Just be sure that if you use Excel, you type =STDEV.S; =VAR.S, and so on, because this is a sample you are calculating something for, NOT a population (.P)

(3) Calculate your test statistic, t, using the formula. (Do this by hand. It actually isn’t that tricky.)

(4) Go to your table and look up your critical t, using a cut-off probability of 0.05 and appropriate degrees of freedom (df). Note that df is just n-1, and is used because the magnitude of the critical t required for significance changes depending on how large the sample size is.

(5) Once you have your critical value from the table, take a look and see how it compares to the one you calculated from the data. Is your calculated t bigger than or smaller than the critical t?

➢ if it is BIGGER, then this tells you that there is enough evidence from your experiment to conclude that a significant difference does indeed exist between the two groups.
➢ if it is SMALLER, then this tells you that there is not enough evidence from your experiment to conclude that a significant difference exist between the two groups.

(6) Now you are ready to look back at your null and alternative hypotheses and make some conclusions. Remember:

➢ if you found a significant difference between your two groups (i.e. calculated test statistic (your data) > critical test statistic (table), then the p-value is < 0.05 (i.e. unlikely to be due to chance)
➢ if you did not find a significant difference between your two groups (i.e. calculated test statistic (your data) < critical test statistic (table), then the p-value is >0.05 (i.e. there is a high probability that these results are simply due to chance)

Based on your p-values, either REJECT THE NULL HYPOTHESIS or FAIL TO REJECT THE NULL HYPOTHESIS. In other words, either consider mean 1 = mean 2, or consider the two means to be different (e.g. mean 1 >/=/≠ mean 2, depending on what your alternative hypothesis originally was.)

(7) Now that you have either rejected or failed to reject your null hypothesis, what does that mean, in the context of your experiment? For example, if you were comparing a new drug to treat high blood pressure to a placebo, and rejected the null hypothesis, you wouldn’t just state in your journal article “we rejected the null hypothesis.” What does that mean? It means that, at the 95% level of confidence (after all, our p-value was < 0.05 if we rejected H₀, so we are at least 95% confident that these results are not just a result of chance), there is sufficient evidence from the data to suggest that the new medication does indeed lead to a significant reduction in blood pressure, compared to the placebo.