P-value

When you perform a hypothesis test in statistics, a p-value helps you determine the significance of your results. Hypothesis tests are used to test the validity of a claim that is made about a population. This claim that's on trial, in essence, is called the null hypothesis.

The alternative hypothesis is the one you would believe if the null hypothesis is concluded to be untrue. The evidence in the trial is your data and the statistics that go along with it. All hypothesis tests ultimately use a p-value to weigh the strength of the evidence (what the data are telling you about the population). The p-value is a number between 0 and 1 and interpreted in the following way:

- A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject the null hypothesis.
- A large p-value (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.
- p-values very close to the cutoff (0.05) are considered to be marginal (could go either way). Always report the p-value so your readers can draw their own conclusions.

For example, suppose a pizza place claims their delivery times are 30 minutes or less on average but you think it's more than that. You conduct a hypothesis test because you believe the null hypothesis, H_o, that the mean delivery time is 30 minutes max, is incorrect. Your alternative hypothesis (H_a) is that the mean time is greater than 30 minutes. You randomly sample some delivery times and run the data through the hypothesis test, and your p-value turns out to be 0.001, which is much less than 0.05. In real terms, there is a probability of 0.001 that you will mistakenly reject the pizza place's claim that their delivery time is less than or equal to 30 minutes. Since typically we are willing to reject the null hypothesis when this probability is less than 0.05, you conclude that the pizza place is wrong; their delivery times are in fact more than 30 minutes on average, and you want to know what they're gonna do about it! (Of course, you could be wrong by having sampled an unusually high number of late pizza deliveries just by chance.)