Optimizely

OPTIPEDIA

Optimization glossary

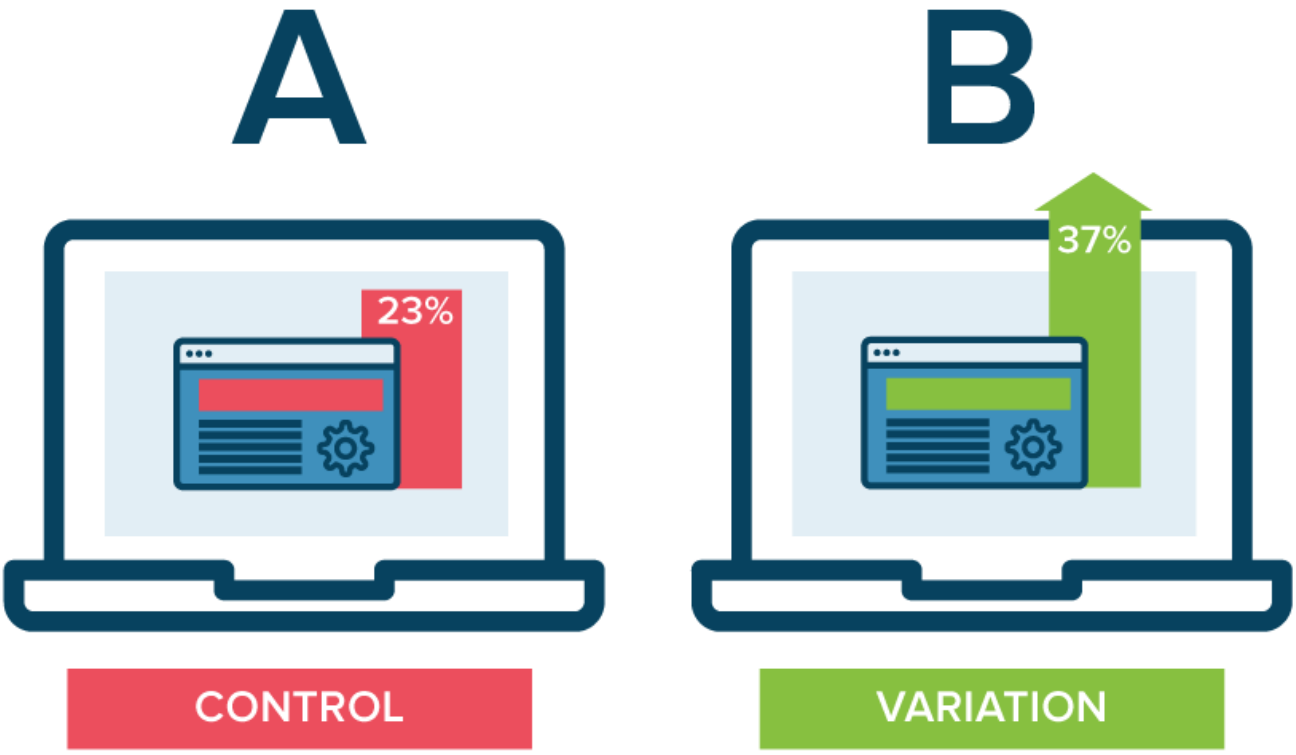
A B C D E F G H I J K L M N O

P Q R S T U V X Y Z

A/B testing

What is A/B testing?

A/B testing (also known as split testing or bucket testing) is a method of comparing two versions of a webpage or app against each other to determine which one performs better. A/B testing is essentially an experiment where two or more variants of a page are shown to users at random, and statistical analysis is used to determine which variation performs better for a given conversion goal.



Running an A/B test that directly compares a variation against a current experience lets you ask focused questions about changes to your website or app and then collect data about the impact of that change.

Testing takes the guesswork out of website optimization and enables data-informed decisions that shift business conversations from "we think" to "we know." By measuring the impact that changes have on your metrics, you can ensure that every change produces positive results.

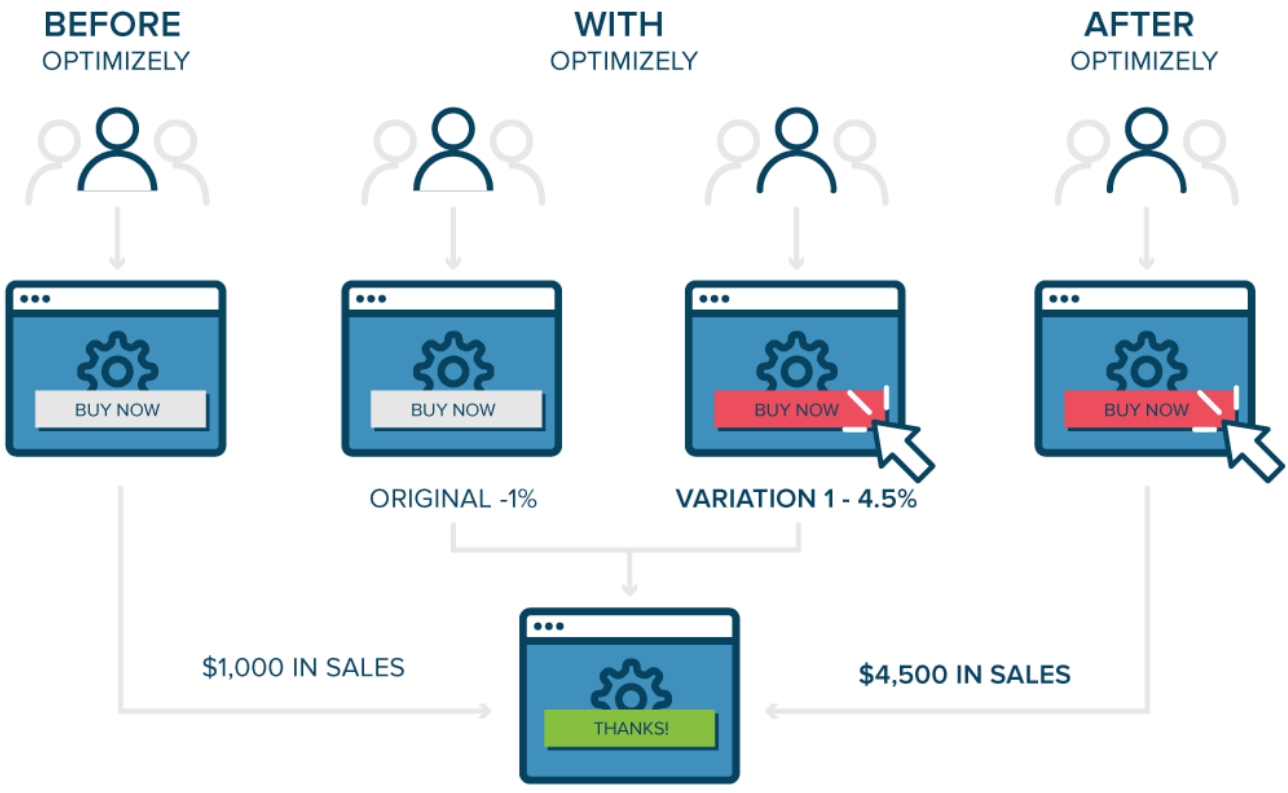
Get started with Optimizely A/B testing

Ready to start your first experiment? Optimizely Web Experimentation can help your set up many types of tests in minutes.

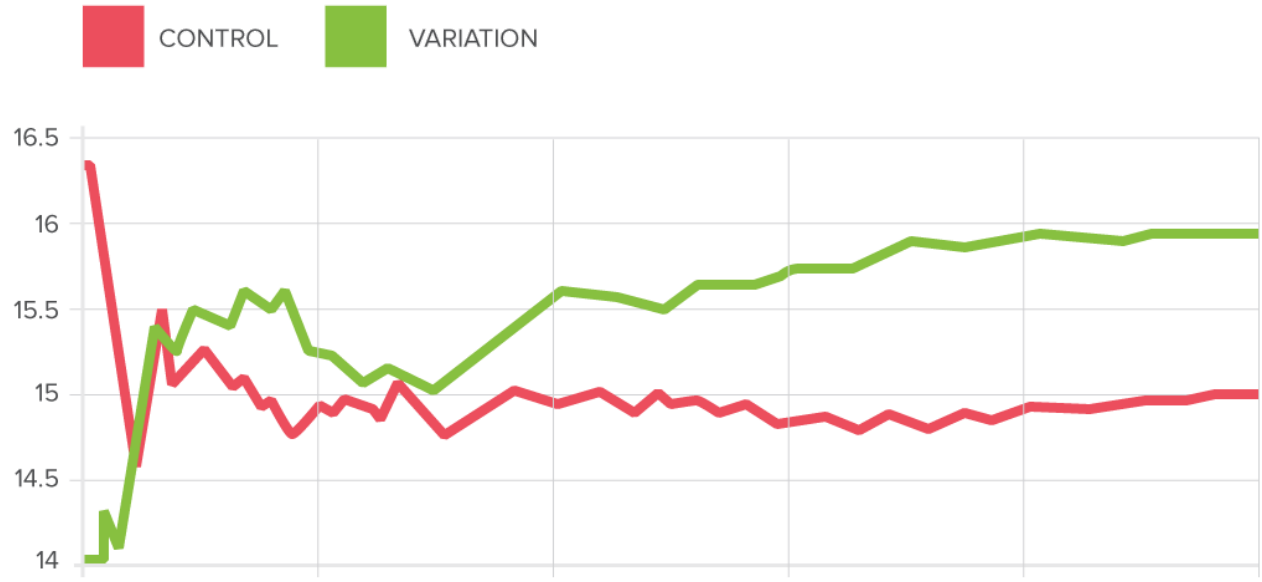
Start A/B testing

How A/B testing works

In an A/B test, you take a webpage or app screen and modify it to create a second version of the same page. This change can be as simple as a single headline, button or be a complete redesign of the page. Then, half of your traffic is shown the original version of the page (known as the control) and half are shown the modified version of the page (the variation).



As visitors are served either the control or variation, their engagement with each experience is measured and collected in a dashboard and analyzed through a statistical engine. You can then determine whether changing the experience had a positive, negative or neutral effect on visitor behavior.



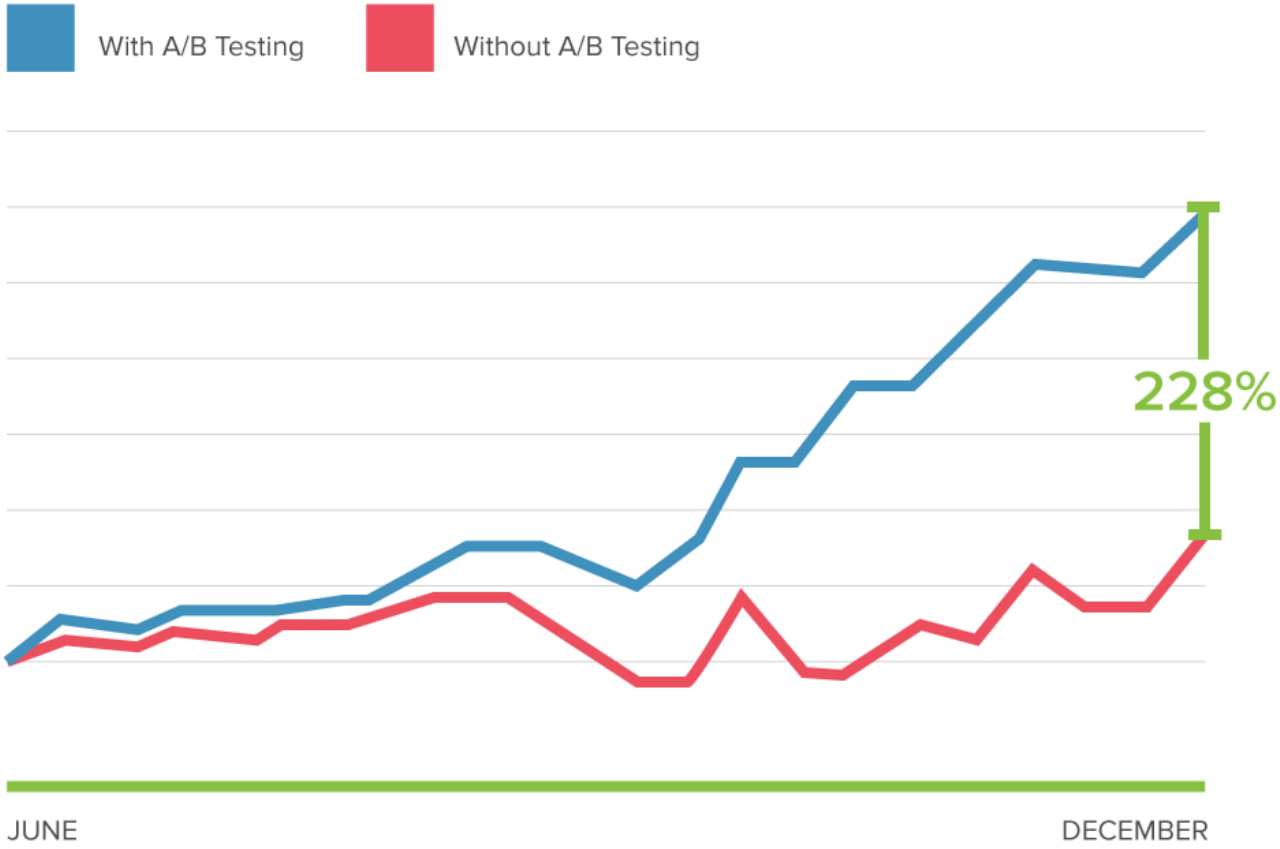
Why you should A/B test

A/B testing allows individuals, teams and companies to make careful changes to their user experiences while collecting data on the results. This allows them to construct hypotheses and to learn why certain elements of their experiences impact user behavior. In another way, they can be proven wrong—their opinion about the best experience for a given goal can be proven wrong through an A/B test.

More than just answering a one-off question or settling a disagreement, A/B testing can be used to continually improve a given experience or improve a single goal like conversion rate over time.

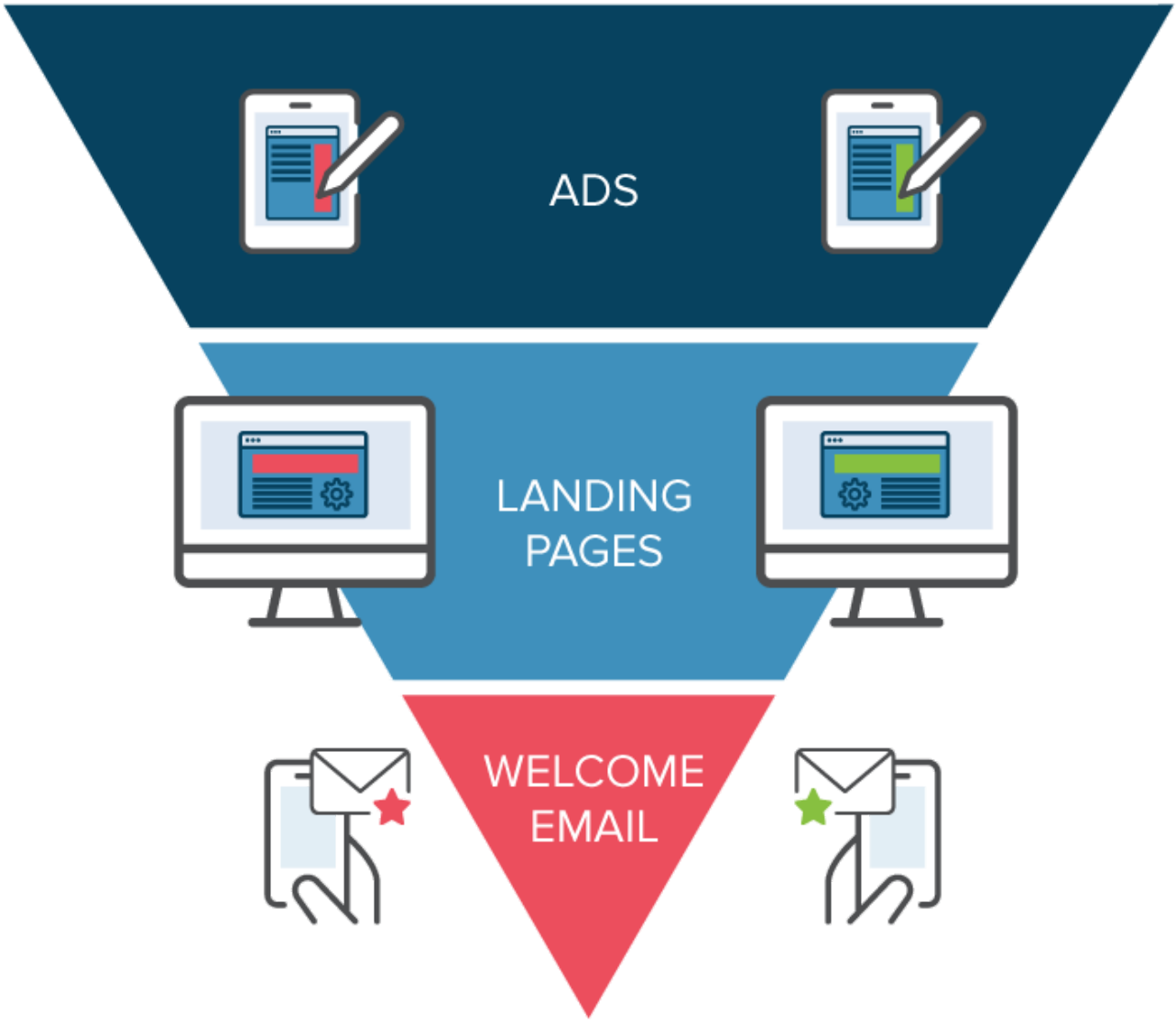
A B2B technology company may want to improve their sales lead quality and volume from campaign landing pages. In order to achieve that goal, the team would try A/B testing changes to the headline, visual imagery, form fields, call to action and overall layout of the page.

Testing one change at a time helps them pinpoint which changes had an effect on visitor behavior, and which ones did not. Over time, they can combine the effect of multiple winning changes from experiments to demonstrate the measurable improvement of a new experience over the old one.



This method of introducing changes to a user experience also allows the experience to be optimized for a desired outcome and can make crucial steps in a marketing campaign more effective.

By testing ad copy, marketers can learn which versions attract more clicks. By testing the subsequent landing page, they can learn which layout converts visitors to customers best. The overall spend on a marketing campaign can actually be decreased if the elements of each step work as efficiently as possible to acquire new customers.



A/B testing can also be used by product developers and designers to demonstrate the impact of new features or changes to a user experience. Product onboarding, user engagement, modals and in-product experiences can all be optimized with A/B testing, as long as goals are clearly defined and you have a clear hypothesis.

A/B testing process

The following is an A/B testing framework you can use to start running tests:

- **Collect data:** Your analytics will often provide insight into where you can begin optimizing. It helps to begin with high traffic areas of your site or app to allow you to gather data faster. Look for pages with low conversion rates or high drop-off rates that can be improved.

- **Identify goals:** Your conversion goals are the metrics that you are using to determine whether or not the variation is more successful than the original version. Goals can be anything from clicking a button or link to product purchases and e-mail signups.

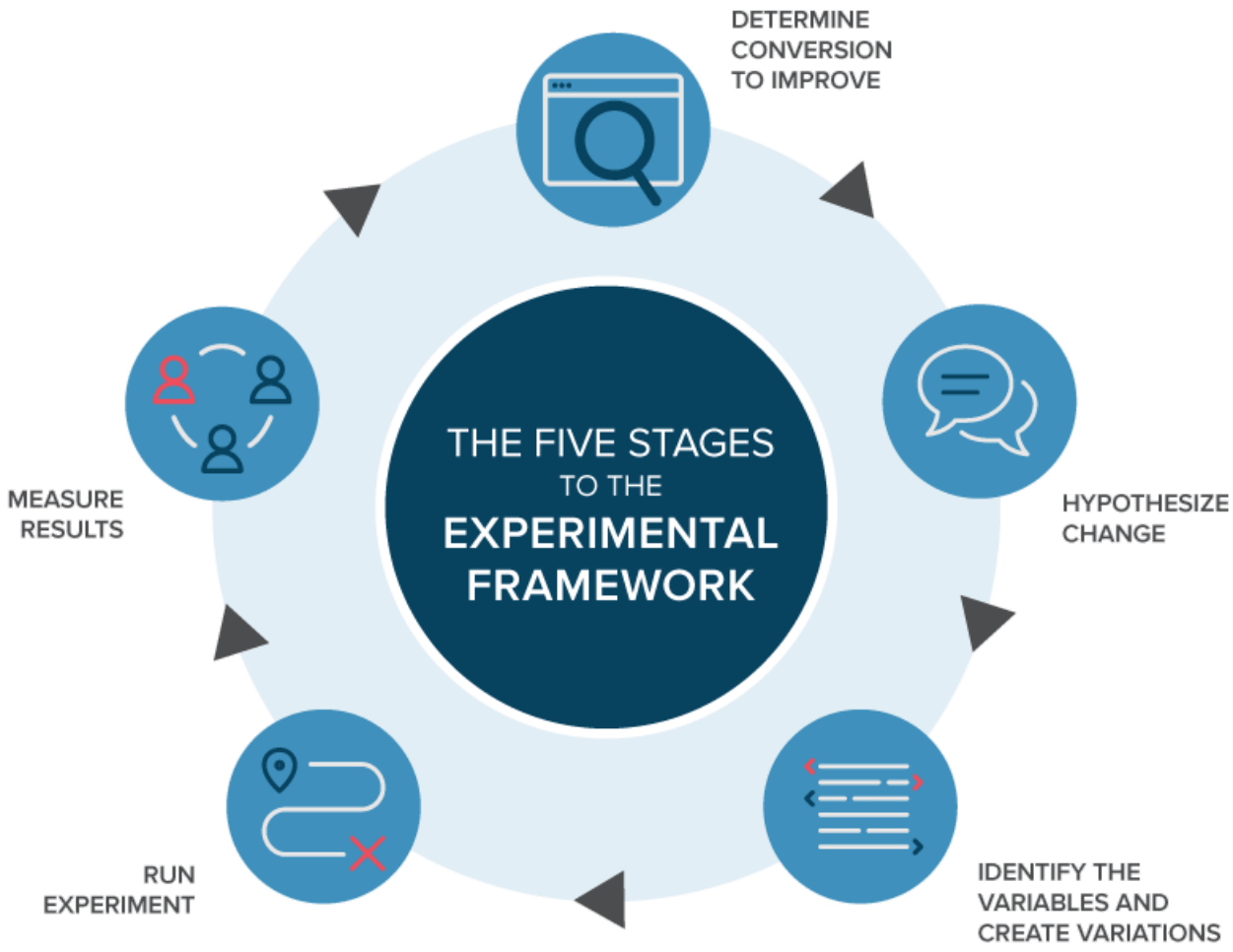
- **Generate hypothesis:** Once you've identified a goal you can begin generating A/B testing ideas and hypotheses for why you think they will be better than the current version. Once you have a list of ideas, prioritize them in terms of expected impact and difficulty of implementation.

- **Create variations:** Using your A/B testing software (like Optimizely), make the desired changes to an element of your website or mobile app experience. This might be changing the color of a button, swapping the order of elements on the page, hiding navigation elements, or something entirely custom. Many leading A/B testing tools have a visual editor that will make these changes easy. Make sure to QA your experiment to make sure it works as expected.

- **Run experiment:** Kick off your experiment and wait for visitors to participate! At this point, visitors to your site or app will be randomly assigned to either the control or variation of your experience. Their interaction with each experience is measured, counted and compared to determine how each performs.

- **Analyze results:** Once your experiment is complete, it's time to analyze the results. Your A/B testing software will present the data from the experiment and show you the difference between how the two versions of your page performed and whether there is a statistically significant difference.

If your variation is a winner, congratulations! See if you can apply learnings from the experiment on other pages of your site and continue iterating on the experiment to improve your results. If your experiment generates a negative result or no result, don't worry. Use the experiment as a learning experience and generate new hypothesis that you can test.



Whatever your experiment's outcome, use your experience to inform future tests and continually iterate on optimizing your app or site's experience.

A/B testing & SEO

Google permits and encourages A/B testing and has stated that performing an A/B or multivariate test poses no inherent risk to your website's search rank. However, it is possible to jeopardize your search rank by abusing an A/B testing tool for purposes such as cloaking. Google has articulated some best practices to ensure that this doesn't happen:

- **No cloaking:** Cloaking is the practice of showing search engines different content than a typical visitor would see. Cloaking can result in your site being demoted or even removed from the search results. To prevent cloaking, do not abuse visitor segmentation to display different content to Googlebot based on user-agent or IP address.

- **Use rel="canonical":** If you run a split test with multiple URLs, you should use the rel="canonical" attribute to point the variations back to the original version of the page. Doing so will help prevent Googlebot from getting confused by multiple versions of the same page.

- **Use 302 redirects instead of 301s:** If you run a test that redirect the original URL to a variation URL, use a 302 (temporary) redirect vs a 301 (permanent) redirect. This tells

search engines such as Google that the redirect is temporary and that they should keep the original URL indexed rather than the test URL.

- **Run experiments only as long as necessary:** Running tests for longer than necessary, especially if you are serving one variation of your page to a large percentage of users, can be seen as an attempt to deceive search engines. Google recommends updating your site and removing all test variations your site as soon as a test concludes and avoid running tests unnecessarily long.

For more information on A/B testing and SEO, see our Knowledge Base article on how A/B testing impacts SEO.

A media company might want to increase readership, increase the amount of time readers spend on their site, and amplify their articles with social sharing. To achieve these goals, they might test variations on:

- Email sign-up modals
- Recommended content
- Social sharing buttons

A travel company may want to increase the number of successful bookings are completed on their website or mobile app, or may want to increase revenue from ancillary purchases. To improve these metrics, they may test variations of:

- Homepage search modals
- Search results page
- Ancillary product presentation

An e-commerce company might want to increase the number of completed checkouts, the average order value, or increase holiday sales. To accomplish this, they may A/B test:

- Homepage promotions
- Navigation elements
- Checkout funnel components

A technology company might want to increase the number of high-quality leads for their sales team, increase the number of free trial users, or attract a specific type of buyer. They might test:

- Lead form components
- Free trial signup flow
- Homepage messaging and call-to-action

A/B Testing Examples

These A/B testing examples show the types of results the world's most innovative companies have seen through A/B testing with Optimizely:

Discovery COMMUNICATIONS

Discovery A/B tested the components of their video player to engage with their TV show 'super fan.' The result? A 6% increase in video engagement.

comScore

ComScore A/B tested logos and testimonials to increase social proof on a product landing page and increased leads generated by 69%.

secret escapes

Secret Escapes tested variations of their mobile signup pages, doubling conversion rates and increasing lifetime value.

Products

Digital Experience Platform

Content Cloud

Commerce Cloud

Intelligence Cloud

Plans & pricing

Resources

Insights

Blog

Events

Developers

Support

Company

Careers

About us
Offices
Executive team
Diversity & inclusion
Press

About working at Optimizely
We're hiring!

Global headquarters

119 5th Ave 7th floor
New York, NY 10003, USA

Contact us

Phone +1 603 594 0249

Language

English

Privacy policy
Website terms of use
Trust center

Optimizely

OPTIPEDIA

Optimization glossary

A B C D E F G H I J K L M N O

P Q R S T U V X Y Z

Statistical significance

What is statistical significance?

Statistical significance is the likelihood that the difference in conversion rates between a given variation and the baseline is not due to random chance.

A result of an experiment is said to have statistical significance, or be statistically significant, if it is likely not caused by chance for a given statistical significance level.

Your statistical significance level reflects your risk tolerance and confidence level. For example, if you run an A/B testing experiment with a significance level of 95%, this means that if you determine a winner, you can be 95% confident that the observed results are real and not an error caused by randomness. It also means that there is a 5% chance that you could be wrong.

Calculate statistical significance

Need to calculate what sample size you'll need to achieve statistical significance? Use our sample size calculator!

Calculate

What does statistical significance really mean?

Statistical significance is a way of mathematically proving that a certain statistic is reliable. When you make decisions based on the results of experiments that you're running, you will want to make sure that a relationship actually exists.

Online web owners, marketers, and advertisers have recently become interested in making sure their A/B test experiments (e.g., conversion rate A/B testing, ad copy changes, email subject line tweaks) get statistical significance before jumping to conclusions.

Testing your hypothesis

Statistical significance is most practically used in statistical hypothesis testing. For example, you want to know whether or not changing the color of a button on your website from red to green will result in more people clicking on it.

If your button is currently red, that's called your "null hypothesis". Turning your button green is known as your "alternative hypothesis". To determine the observed difference in a statistical significance test, you will want to pay attention to two outputs: p-value and confidence interval around effect size.

P-value refers to the probability value of observing an effect from a sample. A p-value of < 0.05 is the conventional threshold for declaring statistical significance.

Confidence interval around effect size refers to the upper and lower bounds of what can happen with your experiment.

Why is statistical significance important for business?

Statistical significance is important because it gives you confidence that the changes you make to your website or app actually have a positive impact on your conversion rate and other metrics. Your metrics and numbers can fluctuate wildly from day to day, and statistical analysis provides a sound mathematical foundation for making business decisions and eliminating false positives.

A statistically significant result isn't attributed to chance and depends on two key variables: **sample size** and **effect size**.

Sample size refers to how large the sample for your experiment is. The larger your sample size, the more confident you can be in the result of the experiment (assuming that it is a randomized sample). If you are running tests on a website, the more traffic your site receives, the sooner you will have a large enough data set to determine if there are statistically significant results. You will run into sampling errors if your sample size is too low.

Effect size refers to the size of the difference in results between the two sample sets and indicates practical significance. If there is a small effect size (say a 0.1% increase in conversion rate) you will need a very large sample size to determine whether that difference is significant

or just due to chance. However, if you observe a very large effect on your numbers, you will be able to validate it with a smaller sample size to a higher degree of confidence.

Beyond these two factors, a key thing to keep in mind is the importance of randomized sampling. If traffic to a website is split evenly between two pages but the sampling isn't random, it can introduce errors due differences in behavior of the sampled population.

For example, if 100 people visit a website and all the men are shown one version of a page and all the women are shown a different version, then a comparison between the two is not possible, even if the traffic is split 50-50, because the difference in demographics could introduce variations in the data. A truly random sample is needed to determine that the result of the experiment is statistically significant.

In the pharmaceutical industry, researchers use statistical test results from clinical trials to evaluate new drugs. Research findings from significance testing indicates drug effectiveness, which can drive investor funding and make or break a product.

Easily calculate statistical significance with Stats Engine

Calculating statistical significance accurately can be a complicated task that requires a solid understanding of statistics and calculus.

Fortunately, you can easily determine the statistical significance of experiments, without any math using stats engine, the advanced statistical model built-in to Optimizely.

Stats Engine operates by combining sequential testing and false discovery rate control signs to deliver statistically significant results regardless of sample size. Updating in real time, Stats Engine will ensure 95% significance level results every time, boosting your confidence in making the right decision for your company and to avoid pitfalls along the way.

To address these common problems, Stats Engine was created to test more in less time. By helping you make statistically sound decisions in real time, Stats Engine adjusts values as needed and shares trustworthy results quickly and accurately.

Start running your tests with Optimizely today and be confident in your decisions.

Digital Experience Platform
Content Cloud
Commerce Cloud
Intelligence Cloud
Plans & pricing

Insights
Blog
Events
Developers
Support

Company

About us
Offices
Executive team
Diversity & inclusion
Press

Careers

About working at Optimizely
We're hiring!

Global headquarters

119 5th Ave 7th floor
New York, NY 10003, USA

Contact us

Phone +1 603 594 0249

Language

English

Privacy policy
Website terms of use
Trust center