

Introduction:-

In marketing and business intelligence, A/B testing is jargon for a randomized experiment with two variants, A and B, which are the control and treatment in the controlled experiment . It is a form of statistical hypothesis testing with two variants leading to the technical term, two-sample hypothesis testing, used in the field of statistics. Other terms used for this method include bucket tests and split testing but these terms have a wider applicability to more than two variants. In online settings, such as web design (especially user experience design), the goal is to identify changes to web pages that increase or maximize an outcome of interest (e.g., click-through rate for a banner advertisement). Formally the current web page is associated with the null hypothesis.

How is A/B Testing performed?

As the name implies, two versions (A and B) are compared, which are identical except for one variation that might affect a user's behavior. Version A might be the currently used version (control), while version B is modified in some respect (treatment). For instance, on an e-commerce website the purchase funnel is typically a good candidate for A/B testing, as even marginal improvements in drop-off rates can represent a significant gain in sales. Significant improvements can sometimes be seen through testing elements like copy text, layouts, images and colors, but not always. The vastly larger group of statistics broadly referred to as Multivariate testing or multinomial testing is similar to A/B testing, but may test more than two different versions at the same time and/or has more controls, etc. Simple A/B tests are not valid for observational, quasi-experimental or other non-experimental situations, as is common with survey data, offline data, and other, more complex phenomena.

An emailing campaign example

A company with a customer database of 2000 people decides to create an email campaign with a discount code in order to generate sales through its website. It creates two versions of the email with different call to action (the part of the copy

which encourages customers to do something — in the case of a sales campaign, make a purchase) and identifying promotional code.

To 1000 people it sends the email with the call to action stating, "Offer ends this Saturday! Use code A1",

and to another 1000 people it sends the email with the call to action stating, "Offer ends soon! Use code B1".

All other elements of the emails' copy and layout are identical. The company then monitors which campaign has the higher success rate by analyzing the use of the promotional codes. The email using the code A1 has a 5% response rate (50 of the 1000 people emailed used the code to buy a product), and the email using the code B1 has a 3% response rate (30 of the recipients used the code to buy a product). The company therefore determines that in this instance, the first Call To Action is more effective and will use it in future sales. A more nuanced approach would involve applying statistical testing to determine if the differences in response rates between A1 and B1 were statistically significant (that is, highly likely that the differences are real, repeatable, and not due to random chance).

In the example above, the purpose of the test is to determine which is the more effective way to encourage customers to make a purchase. If, however, the aim of the test had been to see which email would generate the higher click-rate – that is, the number of people who actually click onto the website after receiving the email – then the results might have been different.

For example, even though more of the customers receiving the code B1 accessed the website, because the Call To Action didn't state the end-date of the promotion many of them may feel no urgency to make an immediate purchase. Consequently, if the purpose of the test had been simply to see which email would bring more traffic to the website, then the email containing code B1 might well have been more successful. An A/B test should have a defined outcome that is measurable such as number of sales made, click-rate conversion, or number of people signing up/registering

Segmentation and targeting

A/B tests most commonly apply the same treatment (e.g., user interface element) with equal probability to all users. However, in some circumstances, responses to treatments may be heterogeneous. That is, while a treatment A might have a higher response rate overall, treatment B may have an even higher response rate within a specific segment of the customer base. For example, the breakdown of the response rates by gender could have been:

	Overall	Men	Women
Total sends	2,000	1,000	1,000
Total responses	80	35	45
Treatment A	50 / 1,000 (5%)	10 / 500 (2%)	40 / 500 (8%)
Treatment B	30 / 1,000 (3%)	25 / 500 (5%)	5 / 500 (1%)

In this case, we can see that while treatment A had a higher response rate overall, treatment B actually had a higher response rate with men.

As a result, the company might select a segmented strategy as a result of the A/B test, sending treatment B to men and treatment A to women in future. In this example, a segmented strategy would yield an increase in expected response rates from 5% – $\frac{40+10}{500+500}$ – to 6.5% – $\frac{40+25}{500+500}$ – constituting a 30% increase.

It is important to note that if segmented results are expected from the A/B test, the test should be properly designed at the outset to be evenly distributed across key customer attributes, such as gender. That is, the test should both (a) contain a representative sample of men vs. women, and (b) assign men and women randomly to each “treatment” (treatment A vs. treatment B). Failure to do so could lead to experiment bias and inaccurate conclusions to be drawn from the test. This segmentation and targeting approach can be further generalized to include multiple customer attributes rather than a single customer attribute – for example, customers' age AND gender – to identify more nuanced patterns that may exist in the test results.

Technical Details

CTR(Click Through Rate) :- Click-through rate (CTR) is the number of users that click on a specific link out of the total users that view a page, email, or advertisement. It is commonly used to measure the success of an [online advertising](#) campaign for a particular website as well as the effectiveness of email campaigns..

APPRISER SDK:- is an SDK being developed by our team which would help developers monitor their app's conversion rate as per different variants.

Visual Editor :- By using a visual editor a developer can create different variants and link every variant to his developer account for testing and monitoring .

Dashboard:- Every developer or a firm would have its dashboard where analytics of different variants would be provided .

CONCLUSION:-

A/B testing has been marketed by some as a change in philosophy and business strategy in certain niches, though the approach is identical to a between-subjects design, which is commonly used in a variety of research traditions. A/B testing as a philosophy of web development brings the field into line with a broader movement toward evidence-based practice. Our product Appriser would help developers keep a track of their user's engagement and behaviour and provide them valuable insights and feedback on how exactly to increase user engagement and retention .

REFERENCES:-

- <https://blog.kissmetrics.com/ab-testing-introduction/>
- <https://vwo.com/ab-testing/>
- <http://www.smashingmagazine.com/2010/06/the-ultimate-guide-to-a-b-testing/>