

Der Einsatz von A/B Tests: Welche Variante ist besser, A oder B?

Der A/B Test (auch split test genannt) ist eine experimentelle Methode, die häufig im Rahmen des Online Marketing eingesetzt wird. In den meisten Fällen geht es dabei um die Ermittlung oder Verbesserung der sog. „Conversion Rate“. Sie bezeichnet das Verhältnis der Anzahl der KäuferInnen eines Produkts zur Anzahl der Personen, die sich über das Produkt informieren. Besuchen z. B. 10.000 Personen die Website eines Produkts und kaufen davon 500 das Produkt, so beträgt die Rate 5%.

Mit dem A/B-Test wird häufig die Auswirkung auf die Conversion Rate ermittelt, die verschiedene Preise, Designs oder Werbemaßnahmen für ein Produkt in einem eShop oder auf einer Website haben. Ist der Preis des Produkts angemessen oder sollte bei der Preisgestaltung etwas verändert werden? Welche Fotos überzeugen potenzielle KäuferInnen am meisten von der Qualität des Produkts? Bemerkten die BesucherInnen die aktuelle Rabattaktion?

ZIELSETZUNG

Um derartige Fragen zu beantworten, bietet es sich an, verschiedene Versionen der Website gegeneinander zu testen. Dafür ist zunächst ein Ziel zu formulieren, das durch eine Veränderung der Website erreicht werden soll. Dies muss nicht in jedem Fall eine Verbesserung der Conversion Rate sein, auch andere Ziele kommen in Frage. Wichtig ist aber in jedem Fall, das Ziel so zu formulieren, dass es messbar ist. Nur so lässt sich beurteilen, ob die Veränderung als Erfolg gewertet werden kann.

HYPOTHESEN UND VORBEREITUNG

Um das Ziel zu erreichen, ist zunächst eine Hypothese darüber zu aufzustellen, durch welche Veränderung der Website dies ermöglicht wird. Da gerade im eCommerce-Bereich vermeintliche Kleinigkeiten bei der Darstellung des Produkts die Conversion Rate maßgeblich beeinflussen können, betrifft die Veränderung meist nur ein einzelnes Detail, wodurch zusätzlich zur ursprünglichen Version A der Website eine neue Variante B entsteht (was die Namensgebung A/B Tests erklärt). So könnte beispielsweise ein Button, der direkt zu einem Produkt führt, dessen Verkauf gesteigert werden soll, farblich hervorgehoben werden, um mehr Aufmerksamkeit zu erregen.

DURCHFÜHRUNG

Zur Durchführung des Vergleichs werden die BesucherInnen der Website für einen festgelegten Zeitraum zufällig in zwei Gruppen „gesplittet“ (siehe Abb. 1). Die sog. Kontrollgruppe wird auf die unveränderte Variante (A) der Website geleitet, die Experimentalgruppe hingegen auf die veränderte Variante (B). Technisch wird die Verteilung der Testpersonen meist über ein Java Script gesteuert. Außerdem werden Cookies verwendet, um Nutzenden, die die Website mehrfach aufsuchen, zu der Variante zu führen, die für sie vorgesehen ist.

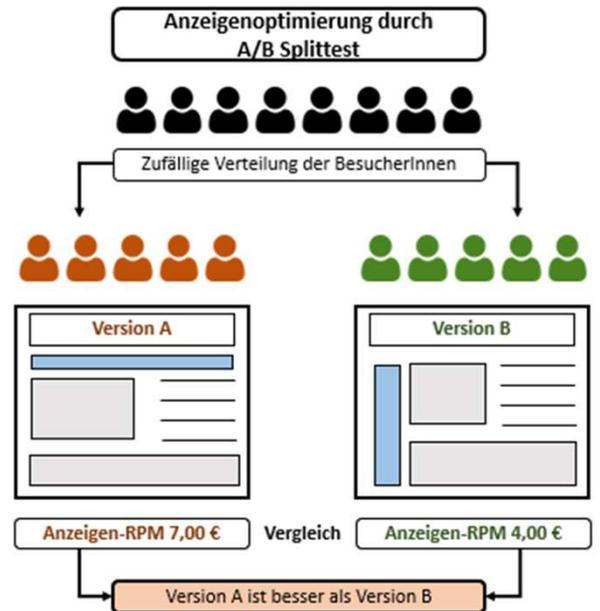


Abbildung 1: Anzeigenoptimierung durch A/B Splittest.

WEITERFÜHRENDE INFOS

Zu A/B Tests:

<https://www.smashingmagazine.com/2010/06/the-ultimate-guide-to-a-b-testing/>

<https://de.wikipedia.org/wiki/A/B-Test>

<https://www.kameleoon.com/de/ab-testing>

Zu statistischen Methoden:

Sedlmeier, P. & Renkewitz, F. (2008). Forschungsmethoden und Statistik in der Psychologie. Pearson Studium: München.

Bildquelle:

https://www.flaticon.com/de/kostenloses-icon/nutzer_1077114?term=nutzer&page=1&position=1&origin=tag&related_id=1077114

VERSUCHSDESIGN UND DATENERFASSUNG

Methodisch betrachtet basiert der A/B Test auf einem einfachen Kontrollgruppen-Design, mit dem untersucht wird, ob sich die beiden Ausprägungen A und B der unabhängigen Variable „Variante“ voneinander unterscheiden. Als abhängige Variable können neben der Conversion Rate je nach Fragestellung auch andere Maße dienen, wie z. B. die Anzahl der BesucherInnen, die einen News Letter abonnieren oder die sich als Nutzende registrierten lassen. Am günstigsten ist es, wenn die benötigten Daten mit gängigen Tools zur Website Statistik, wie z. B. Google Analytics, erhoben werden können. Mittlerweile existiert eine Vielzahl derartiger Tools, die kostengünstig angeboten werden oder sogar kostenfrei zur Verfügung stehen.

AUSWERTUNG UND UMSETZUNG DES ERGEBNISSES

Methodisch betrachtet basiert der A/B Test auf einem inferenzstatistischen Test. Um sicherzustellen, dass etwaige Unterschiede zwischen den Varianten A und B nicht zufällig entstanden sind, werden statistische Verfahren eingesetzt. Handelt es sich um den Vergleich von Häufigkeiten sollte auf parameterfreie Tests, wie beispielsweise den Chi-Quadrat-Test, zurückgegriffen werden. Werden andere Daten erhoben, wie z. B. die Verweildauer auf der Website, kann ein Vergleich der Gruppen mit einem parametrischen Test, wie dem t-Test für unabhängige Stichproben, erfolgen. Ist statistisch ein signifikanter Unterschied ermittelt worden, so steht fest, ob Variante A oder B zu bevorzugen ist. Entsprechend sollte die vorgenommene Veränderung der Website entweder übernommen oder verworfen werden.

DIE GRENZEN VON A/B TESTS

Als Limitation ist zu nennen, dass ein A/B Test – anders als ein Usability-Test – nicht geeignet ist, um Usability-Probleme aufzudecken. Mit dem Test kann zwar die beste Variante identifiziert werden, aber die Gründe für Qualitätsunterschiede lassen sich damit nicht ermitteln. Will man herausfinden, wie die Unterschiede zustande kommen, muss auf andere Verfahren, wie z. B. Lautes Denken, Interviews oder Fokusgruppen, zurückgegriffen werden, die qualitative Daten zu den möglichen Ursachen liefern.

ERWEITERUNG DES A/B TESTS

Soll nicht nur die Veränderung eines Merkmals der Website, wie die Farbe eines Button (z. B. rot versus blau) untersucht werden, sondern eine weitere, wie dessen Positionierung (z. B. am linken Rand, versus mittig, versus am rechten Rand), so lässt sich dies nicht mehr mit einem einfachen A/B Test untersuchen. Im genannten Beispiel entstehen durch die Kombination der 2x3 Merkmale sechs Varianten, die gegeneinander getestet werden müssen. Dies führt nicht nur dazu, dass die technische Realisierung komplizierter wird, sondern auch dazu, dass für die statistische Auswertung anspruchsvollere Verfahren eingesetzt werden müssen.

Der erhöhte Aufwand hat allerdings einen Vorteil. Dadurch, dass die beiden Merkmale nicht in separaten Tests, sondern gemeinsam analysiert werden, erhält man differenziertere Ergebnisse. Testet man Farbe und Position getrennt, würde man zwar erfahren, welche Farbe und welche Position am besten abschneiden, aber nicht, welche Kombination der beiden Merkmale am vorteilhaftesten ist. Dies lässt sich ermitteln, indem multivariate Verfahren, wie z. B. Varianzanalysen, eingesetzt werden, die Aufschluss über die Interaktion von Farbe und Position geben und dadurch zeigen, welche Merkmalskombination am optimalsten ist, um das gesetzte Ziel zu erreichen.

DIE STÄRKEN VON A/B TESTS

Durch das Testen von realen NutzerInnen im laufenden Betrieb der Website erhält man Daten, deren Authentizität im Zusammenspiel mit der hohen Standardisierung der Testdurchführung erheblich zur Reliabilität und Validität der Ergebnisse beiträgt. Außerdem wird durch das Besuchsaufkommen einer Website meist eine Stichprobengröße erreicht, die weit über der von Usability-Tests oder Befragungen liegt. Dadurch können auch kleine Effekte, die eine Veränderung nach sich zieht, erfasst und statistisch abgesichert werden. Des Weiteren ist das Verfahren vergleichsweise kostengünstig, da weder zusätzliche Hardware noch teure Tools angeschafft werden müssen und die Implementierung des Tests nach geringfügigen Anpassungen wieder verwendbar ist. Dies ermöglicht nicht nur eine längerfristige Verwertung zum Testen neuer Varianten, sondern erweitert auch die Einsatzmöglichkeiten des Verfahrens.