# **ClickQuest**: Enhancing User Engagement through **A/B/n Testing** in Library Portal

2024

# 1. Introduction

A/B/n testing, a powerful method for optimizing web page performance, involves comparing multiple variations of a webpage element to determine which one yields the best results. In the context of web design, this testing methodology, often referred to as split testing, enables organizations to make data-driven decisions by analyzing user interactions and behaviors.

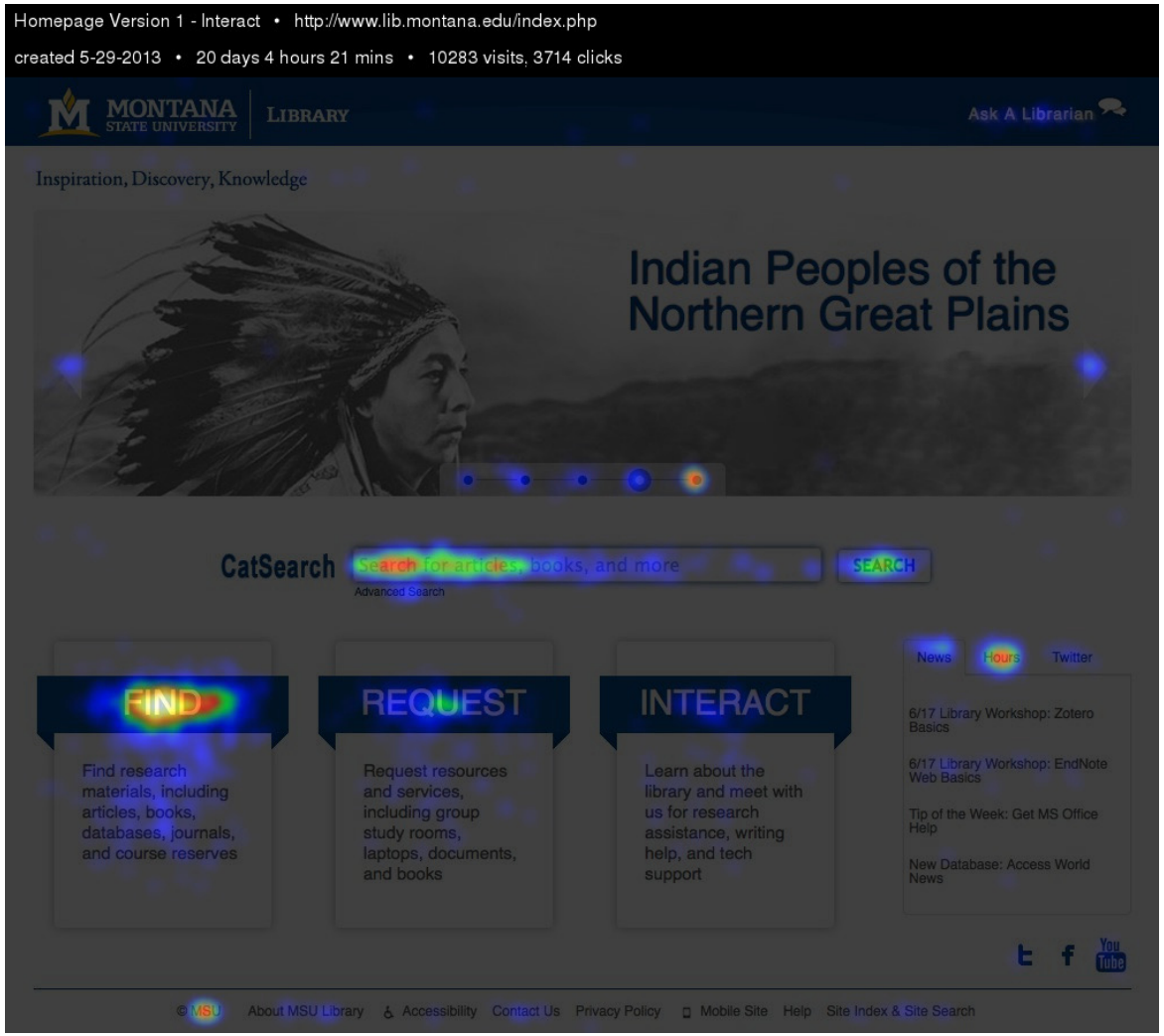
By randomly presenting different versions to users, A/B/n testing allows for the identification of the most effective design, layout, or content, leading to enhanced user experience and increased engagement. This iterative approach empowers teams to refine and tailor web pages for optimal outcomes, making it an invaluable tool for achieving digital success.

Photo by Denys Nevozhai on Unsplash

# Problem Overview

The University of Montana, USA, provides a range of student support services, including a comprehensive library accessible through its web page. However, the library's data analysis revealed a substantial disparity in the click-through rates (CTR) of different page categories. Notably, the "Interact" category exhibited a remarkably low conversion rate of 2%, prompting the IT team to investigate the cause.

Suspecting that the category name might be unclear to users, the team proposed four alternative names: "Connect," "Learn," "Help," and "Service." To validate which name resonates best with students and potentially increase the click-through rate, an A/B/n test was conducted over three weeks, from May 29, 2013, to June 18, 2013. The experiment aimed to ensure equal user exposure to each variation, providing valuable insights to optimize the category and, ultimately, enhance the user experience.



## 1.1 Data Source

The dataset was collected by Google Analytics consists in observations for 5 variations: interact (control), connect, learn, help and services. You can download it here:

https://scholarworks.montana.edu/xmlui/handle/1/3507

## 1.2 Analysis Objective

The project aims to address the following hypothesis:

$H_0$: There is no difference between the CTR of the page variants.

$H_1$: There is a difference between the CTR of the page variants.

# 1.3 Understanding the dataset

- All the variants are listed on the dataset. Interact is the control variant. The other *n*-ones we want to discover what is the best.
- Our strategy are based on clicks, not on the visitors;
- The conversion is the rate below:

$$conversion = \frac{clicks\_link}{clicks\_all}$$

- This data was collected for almost 3 weeks

| | variant | visits | clicks_all | clicks_link | conversion |
|---|---|---|---|---|---|
| **0** | interact | 10283 | 3714 | 42 | 0.011309 |
| **1** | connect | 2742 | 1587 | 53 | 0.033396 |
| **2** | learn | 2747 | 1652 | 21 | 0.012712 |
| **3** | help | 3180 | 1717 | 38 | 0.022132 |
| **4** | services | 2064 | 1348 | 45 | 0.033383 |

# 1.3 Understanding the dataset

- We need to determine the ideal **sample size** to ensure it is statistically significant.
- Furthermore, we are assuming 3 informations:

    1. **Confidence level** = 95% (1- $\alpha$)
    $\alpha$ = 0.05

    2. **Power effect** = 80% (0.80)

    3. **Effect size** (using Chi-squared to discover this)

- We find out that:

| **222** | **1110** |
| --- | --- |
| Minimum sample size per variant | Total sample size |

- Our sample size requirement **is satisfied**, according to the dataset figure on the previous slide.

# 1.4 Methodology

- Now using chi-squared and p-value, we want to discover if there are difference between the contrl variant and the 4 others (connect, learn, help and service)

```
Chi square: 46.33660181942126 - p_value: 2.0959498129984563e-09
```

- The p-value < 0.05 lead us **to Reject $H_0$.** So, there is a significant statistical difference between the control and the other variants. However, we don't know the difference among the those *4* variants.

- A correction of the p-value using the Bonferroni method is used to control the Type I error, which occurs when the null hypothesis is mistakenly rejected. When conducting multiple tests, the probability of committing a Type I error increases and the Bonferroni correction helps offset this increase by making the significance threshold more stringent.
- This means the correction reduces the chance of declaring a difference as significant when, in reality, it is not.
- Correcting the p-value using Bonferroni ins an A/B/n test is important to ensure the results are **statistically robust** when multiple comparisons are being conducted.

# 1.4 Methodology

```
('interact', 'connect'): p_value:5.3676772349808135e-08; corrected_p_value:5.367677234980813e-07; Reject:True
('interact', 'learn'): p_value:0.7616980743361713; corrected_p_value:1.0; Reject:False
('interact', 'help'): p_value:0.0031030587017400212; corrected_p_value:0.03103058701740021; Reject:True
('interact', 'services'): p_value:1.798089447385411e-07; corrected_p_value:1.7980894473854111e-06; Reject:True
('connect', 'learn'): p_value:0.00013292868361715983; corrected_p_value:0.0013292868361715984; Reject:True
('connect', 'help'): p_value:0.06144184057612575; corrected_p_value:0.6144184057612575; Reject:False
('connect', 'services'): p_value:1.0; corrected_p_value:1.0; Reject:False
('learn', 'help'): p_value:0.0508958228881819; corrected_p_value:0.5089582288818191; Reject:False
('learn', 'services'): p_value:0.00020374035733741825; corrected_p_value:0.0020374035733741825; Reject:True
('help', 'services'): p_value:0.07301998638337415; corrected_p_value:0.7301998638337415; Reject:False
```

- The first 4 rows are the interactions among control and each variant;
- The variant *learn* have the **corrected p-value** > 0.05. This means that using *interact* is better;

# 2. Conclusion

Considering the sample size and the result of the p-value corrected by the Bonferroni method, the University of Montana library has the option to use 3 out of the 4 proposed names, **as there will be no difference among the variants** in terms of the number of clicks.

Since there is no significant difference, I would **recommend** using the names guided by the lowest corrected p-value for the composition of the institution's new web page, which are:

Interact X **Connect** ->  corrected p-value = 5.367677234980813e-07

Interact X **Services** -> corrected p-value = 1.7980894473854111e-06

Interact X **Help** -> corrected p-value = 0.031030587017400 21

# Thank you =)

Jonatas Vieira

jonatasv@gmail.com

in/jonatas-vieira/