

DA106 Mastery Project I – A/B Test Recommendation

1 Introduction

An A/B test was conducted to evaluate the effectiveness of a new banner. The primary metrics of interest were the Conversion Rate (CR) and the Average Amount Spent (AAS).

2 Key Findings

Conversion Rate (CR): At a significance level of $\alpha = 0.05$, a large and statistically significant increase in the Conversion Rate due to the banner was observed. Specifically, the Conversion Rate in the control group was 3.92% and 4.63% in the variant group, resulting in an observed mean difference of 0.71%. The 95% CI for this difference is [0.35%, 1.07%]. The z-test yielded a z-statistic of 3.8643 and a $p < 0.001$, indicating that the observed difference is statistically highly significant. The effect size, measured using Cohen's h , is 3.86, indicating that the observed effect is positive and very large. This suggests that the banner has a positive impact on conversions and could attract more or new customers, looking for food and drink offerings especially.

Average Amount Spent (AAS): At the same significance level ($\alpha = 0.05$), the mean Average Amount Spent in the control group was \$3.37, and in the variant group, it was \$3.39, resulting in a minimal observed mean difference of \$0.02. The 95% CI for this difference is [-\$0.44, \$0.47]. The t-test produced a t-statistic of 0.0704 with 24599 degrees of freedom and a p-value of 0.943, indicating that the observed difference is statistically insignificant. The effect size, measured using Cohen's d , is 0.0006, indicating that the observed effect is positive and very small. Importantly, there was no observed negative impact on the Average Amount Spent.

Stable Long-Term Conversion Without Novelty Effect: In the initial stages of the new banner's introduction, a notable difference in conversion rates was observed, with the treatment group showing a 0.73% higher rate than expected within the first two days. This initial spike is likely a result of small sample variability rather than a novelty effect. As the experiment progressed, the conversion rate difference decreased, challenging the novelty effect hypothesis. By day five, the rate had stabilized to a 0.55% difference, aligning more closely with the control group, and then began to rise again. The sustained conversion rate, even after a potential novelty would have worn off, indicates that the banner maintains lasting appeal. By the thirteenth day, the difference in conversion rates had surpassed initial measurements, reaching a 0.79% increase, which points to a long-term positive impact on conversions that extends beyond any initial novelty response.

The power analysis and forecasting suggest however that for a Minimum Detectable Effect (MDE) on CR of at least 10%, the experiment should be run for at least 24 days. For an MDE on AAS of at least 10%, the recommendation extends to 152 days.

3 Business Considerations

Operational Costs: The monetary and operational costs of re-running the experiment are significant. Moreover, the banner has already been designed, and the current data suggests it doesn't negatively impact the metrics.

User Experience: The banner is not an expensive feature in terms of engineering time or overhead. Even a slight improvement, as indicated by the difference in CR, justifies its introduction. Moreover, with the company's growing food and drink offerings, raising awareness through the banner can potentially boost revenue long term and attract new customers.

Practical Implications: The banner's introduction seems promising, especially if the potential of drawing more users back for subsequent purchases or attracting new customers due to prominently presenting food and drinks offerings is taken into consideration.

4 Recommendation

Given the significant increase in the Conversion Rate, the absence of a negative impact on the Average Amount Spent, combined with the business considerations:

We recommend launching the banner. By doing so, we can continue learning about user interactions with the feature in a real-world setting. If, in the future, metrics suggest any unforeseen negative implications (e. g. on the Average Amount Spent), we can reassess and make necessary adjustments. However, the current data indicates that the banner is poised to offer value to both the users and the business.