

GloBox - A/B Test

Final Report

Maxwell Acha
28/08/2023

Table of Content

1. Executive Summary	2
2. Context	2
Experimental Context:	2
The A/B Test:	2
Dataset Overview:	3
3. Results	3
Conversion Rate:	3
Average Amount Spent:	4
Novelty Effect:	4
Total Amount Spent per day:	5
Power Analysis:	5
4. Recommendations	6
5. Appendix	6

1. Executive Summary

In this analysis, I have examined the result of an A/B test to assess the impact of an additional banner (Treatment group) compared to the existing homepage layout (Control group). Our findings indicate that the treatment group resulted in a statistically significant increase in conversion rates, with a p-value of 0.0001. Considering the low cost of launching the feature, I recommend implementing the new banner to capitalize on this improvement in conversion rates.

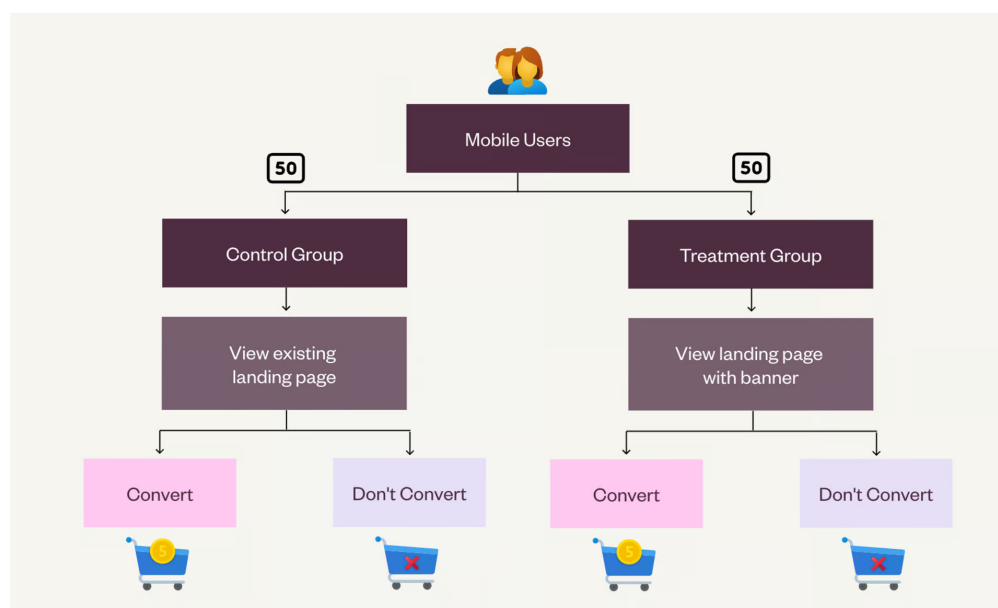
2. Context

Experimental Context:

GloBox is known for offering boutique fashion items and high-end decor products, however in the last period the company have seen a significant growth in the food and drinks products. The motivation behind this A/B test was to bring awareness to the food and drinks category to increase revenue by showing an additional banner on the homepage to half of the users to find out the effect on conversion rate and amount spent (Key Metrics).

The A/B Test:

The experiment was conducted over a span of two weeks (Jan 25th to Feb 6th) and it only affected the mobile website. The users were split randomly and equally between the Control and the Treatment group.



Dataset Overview:

Data were stored into a relational database made up of 3 tables: users (user demographic information), groups (user A/B test group assignment) and activity (user purchase activity, containing 1 row per day that a user made a purchase).

Exploring the database:

- The total number of users who took part in the experiment were 48,943 of which 24343 assigned to the Control group and 24600 to the Treatment group.
- Conversion Rate for all user was 4.28% (Control group 3.92% and Treatment Group: 4.63%)
- The Average amount spent for the Control group was 3.37 and for the Treatment group was 3.39
- 139 user made more than one purchase during the experiment.

The Dataset was explored using PostgreSQL language, more detail about the queries can be found at this [here](#).

3. Results

To determine whether there was a difference in the conversion rate per user between the two groups, I ran a hypothesis test.

Conversion Rate:

H0: There's no difference in conversion rates between Control and Treatment group.

H1: There is a difference between the conversion rate of the two groups.

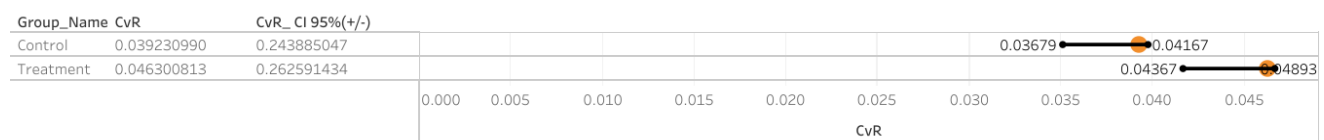
α : 0.05

P-value: 0.0001

Lower Bound: 0.0035

Upper Bound: 0.0107

With p-value = 0.0001, statistically ***significant***. We ***reject*** the null hypothesis that there is no difference in the user conversion rate between the control and treatment. The 95% confidence interval for the parameter lies between 0.0035 and 0.0107, suggesting that we can be confident that the true value falls within this range.



Average Amount Spent:

H0: There's no difference in AVG Amount Spent per user between Control and Treatment group.

H1: There is a difference between the AVG Amount Spent per user of the two groups.

α : 0.05

P-value: 0.944

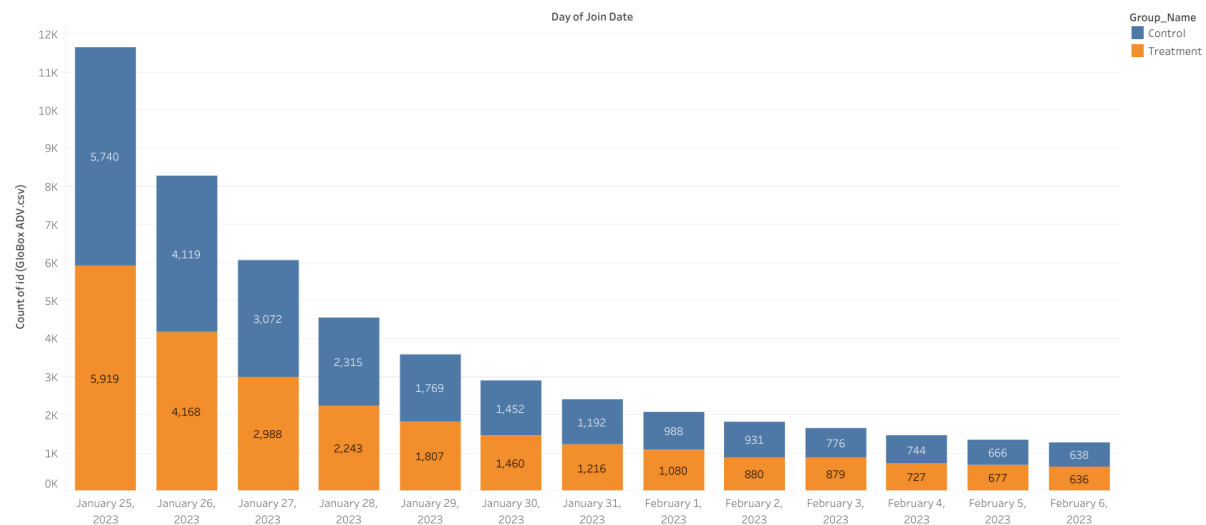
Lower Bound: -0.439

Upper Bound: 0.471

With p-value = 0.944, statistically ***insignificant***. We ***fail to reject*** the null hypothesis that there is no difference in the AVG amount spent per user between the control and treatment. The 95% confidence interval for the parameter falls between -0.439 and 0.471, suggesting that we can be confident that the true value of the parameter lies within this range. Note, it includes the Null Hypothesis value (0).

Novelty Effect:

Total User Joining per Day

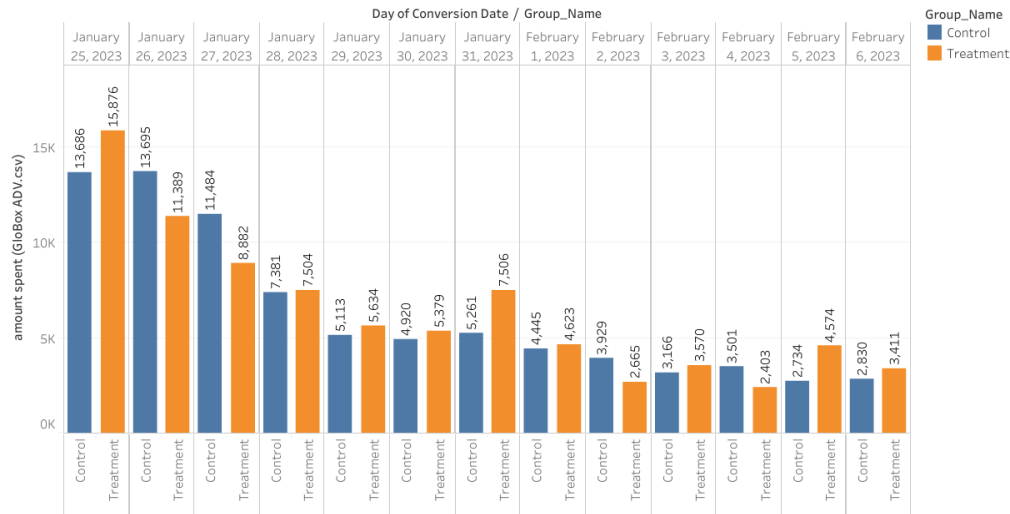


The graph shows that the distribution of the number of users visiting the website each day is equally split between the two group. Therefore, we can assume that there is no novelty effect affecting the experiment

Total Amount Spent per day:

The graph below shows the total amount spent per each day of the experiment. At the end of the experiment, the Treatment group registered a tot amount spend higher than the Control group (difference of \$1,269)

Total Amount Spent per Day



Power Analysis:

Using a 60% Statistical Power with 10% minimum detectable effect and the 3.92% baseline conversion rate of the Control group, the ideal sample size is 24,144 per group. We can then assume that our sample size is normally distributed between the 2 groups.

Question: How many subjects are needed for an A/B test?

Baseline conversion rate:
 %
[link]

Minimum Detectable Effect:
 %
3.53% – 4.31%

☐ Absolute
☒ Relative

The Minimum Detectable Effect is the smallest effect that will be detected (1-β)% of the time.

Conversion rates in the gray area will not be distinguishable from the baseline.

Sample size:

24,144

per variation

Statistical power 1-β: 60% Percent of the time the minimum effect size will be detected, assuming it exists

Significance level α: 5% Percent of the time a difference will be detected, assuming one does NOT exist

4. Recommendations

I highly recommend launching the new banner for the following reasons:

1. Hypothesis testing for the difference between the conversion rate of the two groups resulted in a p-value of 0.0001 which shows statistically significant evidence that there is a difference between the two groups. Further investigations show that the Conversion rate is significantly higher in the Treatment Group.
2. Although the hypothesis test for the AVG amount spent per user between the Control and Treatment group resulted in a p-value = 0.944, which is statistically insignificant. But has we can see the total amount spent at the end of the experiment is higher in the treatment group compared to the control group
3. The cost of implementing the new banner is low.

5. Appendix

1. [SQL queries](#)
2. [Analysis in Spreadsheet](#)
3. [Tableau Workbook](#)