**GlowBox A/B testing**

Matej Papac, 25.11.2023

**Summary :**

We were conducting A/B testing to compare a control group (A) with a treatment group (B) to identify any significant differences.

The Growth team is running an A/B test with a product banner on the website. The control group doesn't see the banner.

Tests showed us a significant difference in conversion rates with a p-value of 0.0001 and

average spent per group was p-value of 0.94385, indicating no significant difference

The power analysis indicated the need for more sample data to proceed; specifically, 70,400 samples are required.

My opinion is to gather more data to continue the analysis effectively.

**Context :**

These are the key values that were utilized in my analysis:

Group A (Control Group): 24,343 participants

Group B (Treatment Group): 24,600 participants

**Conversion Rates:**

Group A (Control Group): 3.92%

Group B (Treatment Group): 4.63%

**Average Amount Spent(dollars):**

Group A (Control Group): 3.37

Group B (Treatment Group): 3.39

A screenshot of a calculator

Description automatically generatedA screenshot of a computer

Description automatically generated

The Z-test for conversion rates showed a significant increase with a p-value of 0.0001, and the confidence interval for the difference in proportions ranged between 0.0035 and 0.0106.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated However, the t-test for the average amount spent did not reveal a significant difference, as indicated by a high p-value of 0.94385, with a confidence interval for the difference in means stretching from -0.4386 to 0.4713.

**A graph with blue and orange lines

Description automatically generated**

The graph shows conversion rates for two groups over two weeks. Group B consistently outperforms Group A, with a sharp increase in the final days, indicating a possible novelty effect from the treatment. However, the available data is insufficient to conclusively determine if this upward trend will persist.

A screenshot of a computer

Description automatically generated

The power analysis reveals that to detect a 10% effect at a 4.27% baseline conversion rate, a sample size of 70.4k is needed, split equally between the test and control groups.

**Results:**

In this section, I will also present analytical insights that will shed light on user engagement and behavioral patterns, complementing the findings from the A/B testing.

A graph of different colored squares

Description automatically generated

Females have higher conversion rates than males in both groups, and the unknown gender category has a moderate conversion rate.

**A graph with numbers and a bar

Description automatically generated**

This histogram shows the distribution of total spending per user in each group. It helps identify spending patterns and the concentration of spending amounts across users within each group.

A screenshot of a computer

Description automatically generated

This indicates the confidence interval for the difference in the average amount spent per user between the two groups. The point estimate suggests a slight increase in spending for the treatment group over the control group, but the confidence interval crosses zero, implying this result is not statistically significant.

A graph with a line and a dot

Description automatically generated with medium confidence

This graph illustrates the confidence interval for the difference in conversion rates between the two groups. The point estimate of 0.71% indicates a higher conversion rate in the treatment group compared to the control group, and the confidence interval does not include zero, suggesting this is a significant difference.

A graph with numbers and a number of items

Description automatically generated with medium confidence

This scatter plot compares the total amount spent by users on different device types within each test group. It helps to identify which devices are contributing more to total spending and may indicate device preference for transactions among users in the test and control groups.

A computer screen shot of text

Description automatically generated

My data cleaning proccess.

**Recommendation:**

Based on our tests a difference in conversion rates is shown, but only a slight variation in spending.   
  
It is advisable to collect more data. This additional data will provide a more robust understanding of customer behavior and inform better decision-making.

**Appendix:**

**https://public.tableau.com/app/profile/matej.papac/vizzes**