A/B Test Analysis Report:

Globox food and Drink Banner Experiment

Submitted by: Rutuja Kale

Index

1. Summary
2. Data and Approach
3. Context
4. Results
5. Recommendation
6. Appendix

Summary

An A/B test was carried out on a sample of Globox customers, during which participants in the treatment group were introduced to a banner advertising the expanding range of food and drinks on offer. These results were collected, and statistical analysis was conducted to determine whether there was a difference between average spend and conversion rate between the control and treatment groups. The results of the analysis suggest that while there was a significant difference between conversion rates, there was no significant difference between the average spend of the two groups. The final recommendation is to launch the banner to all users, regardless of the lack of difference in average spend.

However, its important to continue monitoring the average amount spent and consider conducting further experiments or increasing the sample size to obtain more robust conclusions regarding differences in means. This will provide dipper understanding of the impact of the banner on user spending behaviour and inform future decision making.

Data and Approach

In this section, I outline the Approach and Methodology I used to analyse the A/B test results.

Step 1. Data Extraction

The first step involved extracting the necessary data for the A/B test from our e-commerce platforms database with the help of SQL. This database was consisted of three tables in it:

* Activity: This table contains the user’s purchase activity which containing 1 row per day that a user made a purchase. This table contains the user’s ID, purchase date, device type and the amount spent in USD.
* Groups: This table is about the user’s A/B test group assignment. It includes the user ID, the user’s test group, the date the user joined the test and the device the user visited the page on (I=iOS, A=Android)
* Users: This table is about the user demographic information. It includes the user ID, ISO 3166 alpha-3 country code, the users gender (M = male, F = female, O = other)

Using these three tables I proceeded with SQL queries to gather relevant information, including user interactions, conversion rates and the average amount spent for both the control and treatment groups.

Step 2. Hypothesis Testing with spreadsheet:

Once the data was collected, I preformed the Hypothesis test to compare the conversion rates and Average amount spent between the Group A (Control) and Group B (Treatment). To test the conversion rates and Average amount spent I formulated the following terms:

* Null Hypothesis (H₀): it says, there is NO difference in conversion rates and average amount spent between the two groups.
* Alternative Hypothesis (H₁): it says, there is difference in conversion rates and average amount spent between two groups.

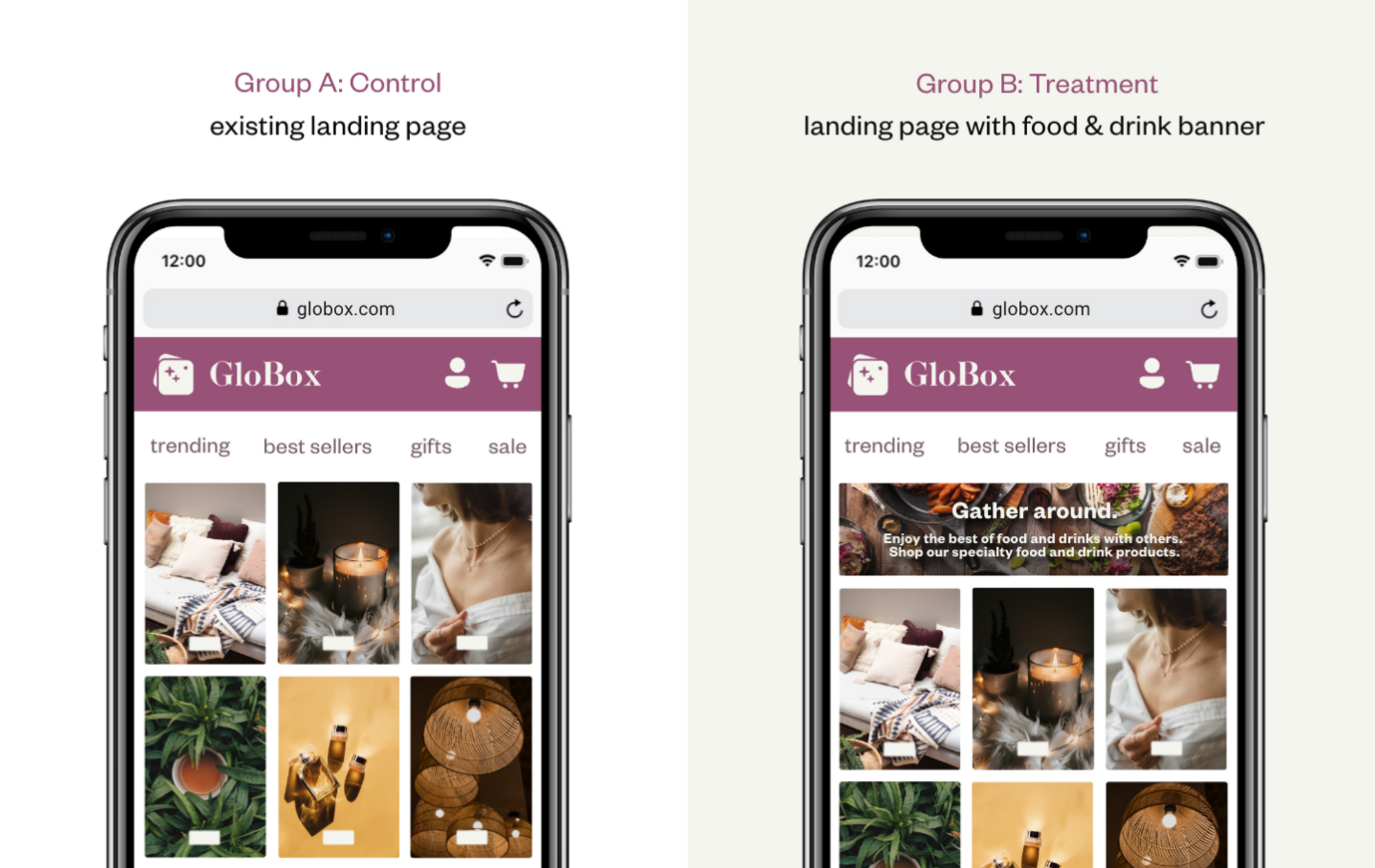
I conducted both the t-test and z- test at a 5% significance level using excel spreadsheets to evaluate the statistical significance of the results.

Step 3. Data visualization with Tableau:

Finally, I used the Tableau to visualize the data to gain deeper understanding into user behaviour and spending patterns. I created the visualizations to analyse the distribution of average amount spent, gender wise analysis, device wise analysis, country and region analysis.

Context

The motivation for the project was to test whether a banner advertising the key food and drinks product at the top of the landing page for the Globox website, would increase revenue, and draw attention to these products to encourage customers to purchase them. To do this A/B test was conducted between the 25th of January 2023 and the 6th of February 2023, in which participant were placed into either the Control group or Treatment Group. The Control group experienced the website in its usual form, and their interactions were collected, alternatively the Treatment group were introduced to the new banner, and their interactions were also collected, and then compared to the Control group .



Analysing the dataset allowed for an investigation into the impact of the food and drink category banner on conversion rates and the average amount spent per user. Additionally, the dataset provided an opportunity to explore user behaviour based on factors such as gender wise, device type, and country and region, offering insights for targeted marketing strategies.

Results

Inferential statistics:

|  |  |  |  |
| --- | --- | --- | --- |
| Group | Sample size | Conversion rate | Average amount spent |
| Control (A) | 24,343 | 3.92% | $3.37 |
| Treatment (B) | 24,600 | 4.63% | $3.39 |

Conversion rate:

The **Null Hypothesis (H₀)** stated that there is no disparity in the conversion rate between Group A and Group B, while the **Alternative Hypothesis (H₁)** stated that there is indeed a difference in the conversion rate between Group A and Group B.A significance level of 0.05 was chosen, and the calculated p-value was 0.0001. The 95% confidence interval for the difference in conversion rates was determined to be 0.0035 to 0.0107.

Based on the statistical analysis, the **null hypothesis was rejected**, indicating a significant difference in conversion rates between both groups. The confidence interval's positive values suggest that the treatment group likely exhibits a higher conversion rate compared to the control group. However, it's essential to interpret the confidence interval correctly. The 95% confidence interval means that if we were to repeat this experiment 100 times, we would expect the true difference in conversion rates to be within 0.35% and 1%. in approximately 95 of those experiments.

Average amount spent:

The Null hypothesis (H₀)stated that there is no difference in the average amount spent per user between the group A(control) and group B(Treatment), while the alternative hypothesis (H₁) stated that there is indeed a difference in the average amount spent between group A( control) and group B(treatment).

The p-value obtained from the t-test was 0.944, and the confidence interval was -0.439 to 0.471.

Based on the statistical analysis, we fail to reject the Null hypothesis, indicating no statistical difference in the average amount spent per user between both groups. The confidence interval containing zero further supports this finding.

1.Distribution of average amount spent:

This graph shows the pattern of average amount distribution per user for each group. The following bar chart represent the count of users on y-axis and Average amount spent on x-axis.

* The majority of the Control Group spent between $40 to $50.
* The majority of Treatment Group spent between $30 to $40.

A graph of a bar chart

Description automatically generated with medium confidence

2. Device wise Analysis

A screenshot of a graph

Description automatically generated A screenshot of a graph

Description automatically generated

* The provided data represents the distribution of device types across two groups: Control (Group A) and Treatment (Group B). The banner had a positive impact on **iOS**users, with the Treatment Group experiencing an uplift of approximately 10.6% in conversion rate among iOS users upon exposure to the banner. It had a slight impact on the average amount spent among iOS users, a decrease of approximately 3% in the average amount spent among iOS users upon exposure to the banner.
* On **Android**devices, the average amounts spent, and conversion rates were lower than on iOS devices, but the Treatment Group indicates a notable uplift of approximately 27% in conversion rate among Android users upon exposure to the banner. It had a positive impact on the average amount spent among Android users, with the Treatment Group showing a moderate increase of 7% in the average amount spent among Android users upon exposure to the banner.

3. Gender wise analysis: