**A/B Test Analysis Report**

**GloBox food and Drink Banner Experiment**

*Presented By: Rutuja kale .* *Date: 24-Dec-2023*

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.

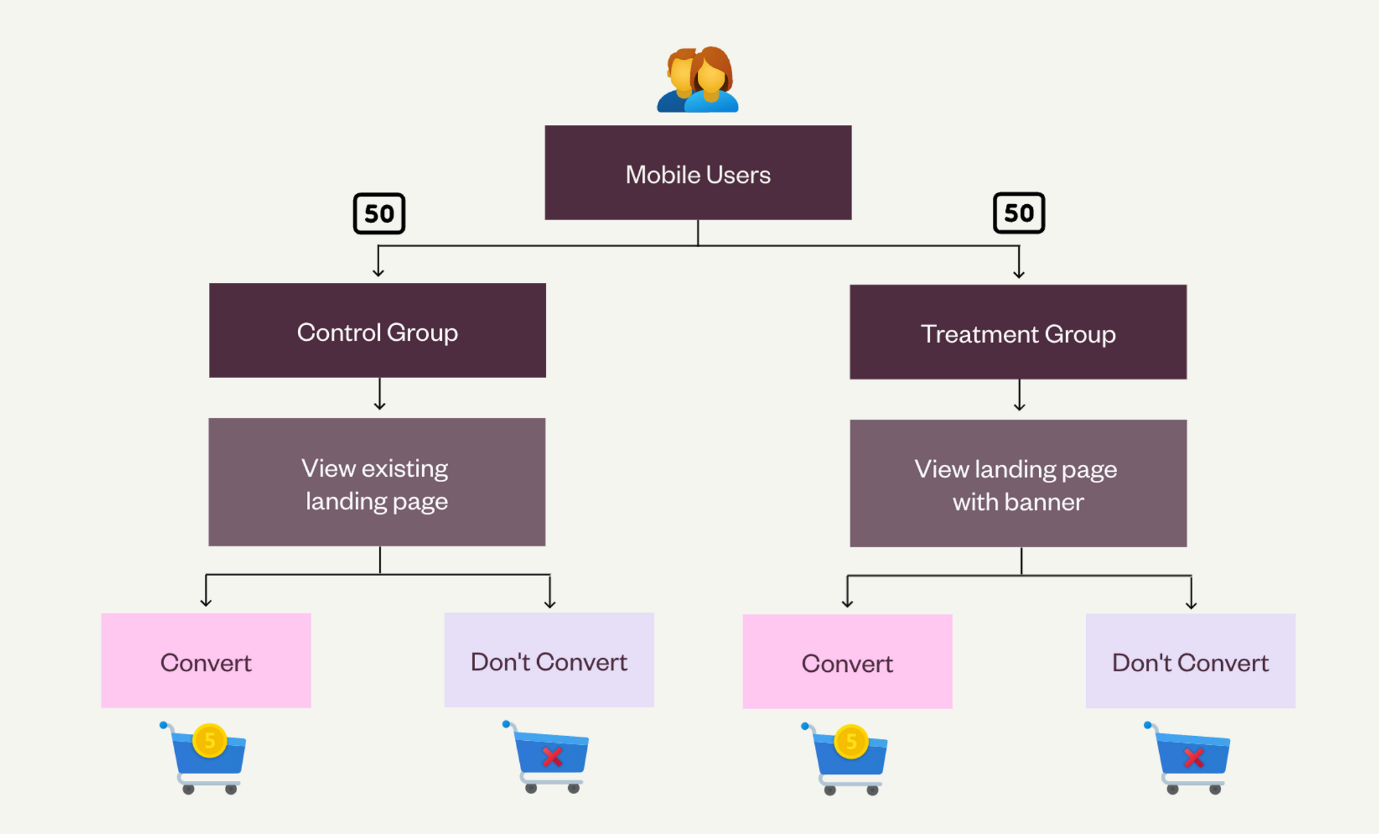
**Step 3. Data visualization with Tableau**:

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.

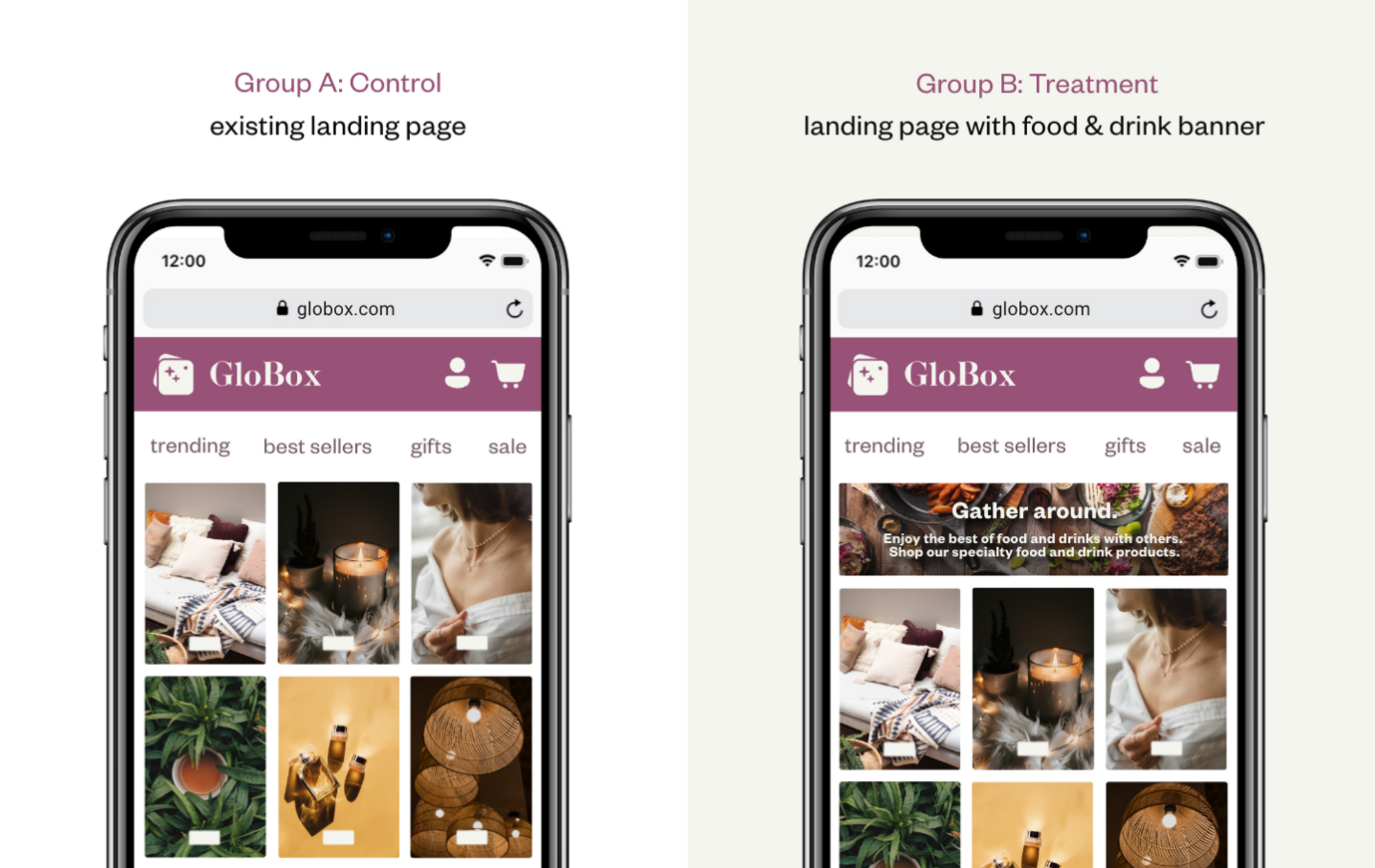
Based on the statistical analysis, the **null hypothesis was rejected**, indicating a significant difference in conversion rates between both groups. The confidence interval's positive values suggest that the treatment group likely exhibits a higher conversion rate compared to the control group. However, 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. based on the graph it is observed that:

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

A screenshot of a computer

Description automatically generated A graph of a number of bars

Description automatically generated

1. **Device wise Analysis**

* 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.

A graph of a comparison of a comparison of a device

Description automatically generated with medium confidence A screenshot of a graph

Description automatically generated

1. **Gender wise analysis:**

Based on the evident, it is observed that,

* Average Amount Spent:
* Control group:
* Male: Males have the lowest Average amount spent among all others.
* Female: Females have highest Average amount spent among all others.
* Treatment group:
  + Male: In the treatment group, Males shows some improvement in Average amount spent. It shows improvement of 15%.
  + Female: In the Treatment group, Female shows decrease of approx. 7%.
* Conversion Rate:
* Control Group:
  + Male: Conversion rate of Males in Control group is lowest among all others.
  + Female: Conversion rate of females in control group is the highest among others.
* Treatment Group:
  + Male: It shows some improvement compared to Control group of 44%.
  + Female: females already had relatively higher conversion rate in Control group. Females also shows improvement of 5% in Treatment group.

A screenshot of a computer

Description automatically generated

A graph with blue and orange squares

Description automatically generated

A blue and orange rectangular boxes

Description automatically generated

1. **Country wise Analysis:**

Based on evidence, it shows that United states and Canada shows highest Conversion rate relatively in Control group among all other countries, While Australia shows the lowest conversion rate in control group. On the other hand, banner shows the postivie impact on canada and United States, it shows respectively 38% and 12% improvements in treatment group.

Based on evidence, it shows that United States and Turkey shows highest average amount spent in control group among all other countries. Also Australia shows the lowest average amount spent in control group. On the other hand, United Kingdom shows the highest average amount spent by 113%. Also Turkey shows the 32% decline in average amount spent in treatment group.

A graph with numbers and a bar chart

Description automatically generated with medium confidence

A graph with numbers and a number of percentages

Description automatically generated

A map of the world

Description automatically generated

* **Novelty Effect:**
  + **Distribution of Average Amount Spent:**

To determine whether there is a novelty effect in the experiment, we can analyze the data provided.

**1. Average Amount Spent by All Users (Control Group vs. Treatment Group):**

Looking at the average amount spent by all users in both groups, we can observe some variations. However, it's essential to consider the overall trends and not focus on individual days. Let's visualize the data to better understand the patterns:

A graph with blue and orange lines

Description automatically generated

From the graph, we can see that there is no clear and consistent trend indicating a significant novelty effect. While there are some fluctuations, the differences between the Control and Treatment groups are not consistently large enough to suggest a novelty effect.

**2. Conversion Rate by All Users (Control Group vs. Treatment Group):**

let's examine the conversion rates for all users in both groups over time. Like the average amount spent, the conversion rates do not exhibit consistent and significant differences between the Control and Treatment groups. Although there are minor fluctuations, there is no indication of a novelty effect affecting the conversion rates.

A graph of a graph with blue and orange lines

Description automatically generated

**3. Average Amount Spent by Converted Users (Control Group vs. Treatment Group):**

Now, let's focus on the average amount spent by only the converted users in both groups.

A graph of a graph showing the average amount of a number of people

Description automatically generated with medium confidence

**4. Time to Purchase:**

Lastly, let's analyze the time it takes for users in both groups to purchase after joining the experience:

A graph with numbers and a bar

Description automatically generated

* **Power Analysis:**

An online power analysis calculator was utilized to determine the appropriate sample size for a study aiming to detect a minimum detectable effect with the desired statistical power. The study focused on comparing the conversion rates of a treatment group to a control group, with the control group's conversion rate of 3.92% considered as the baseline and a minimum detectable effect of 10%.

The power analysis yielded a recommended total sample size of 60,600. (60.6K) This sample size is expected to provide adequate statistical power for detecting the desired effect size.

A screenshot of a computer

Description automatically generated

Additionally, another analysis was performed, which involved comparing two independent means. For this analysis, the difference between the means was set as 10% of the control group's mean, resulting in a value of 0.337. The expected standard deviation used for this analysis was 25.67 units.

A screenshot of a computer

Description automatically generated

Based on these inputs, the power analysis determined that a total sample size of 182,164 would require. This sample size was calculated to achieve a power of 80% and a significance level of 5% (two-sided) for detecting a true difference in means of 0.337 units, assuming a pooled standard deviation of 25.67 units.

The power analysis results indicate that we had a sufficient sample size to confidently evaluate the impact on conversion rates. The effect size for the conversion rate was much larger (18%) than the minimum detectable effect of 10%, and the power analysis confirmed that the sample size used in the A/B test was adequate to detect this effect.

**Recommendation:**

Based on the analysis of the A/B test results and the additional insights provided, there is sufficient information and statistical evidence to make a confident recommendation to launch the food and drink category banner. The A/B test results demonstrated a significant difference in conversion rates between the control group and the treatment group, indicating that the banner has a positive impact on driving conversions.

I recommend launching the food and drink category banner while simultaneously conducting a more detailed analysis of user spending behaviour. This will allow us to better understand the impact of the banner on different customer segments and tailor marketing strategies accordingly.

**Appendix:**

1. [SQL Queries](https://github.com/rtk0422/GloBox/blob/main/SQL%20Queries.sql)
2. [Excel Spreadsheet: Hypothesis Test](https://docs.google.com/spreadsheets/d/1RtyQtPmG86hlGE8WIqSQQwOnsUghku8kKnChQmyrbbU/edit?usp=sharing)
3. [Tableau Visualisation](https://public.tableau.com/views/Globox_17032871197480/Conversion_rate?:language=en-GB&:display_count=n&:origin=viz_share_link)