

Globox New Banner

A/B Test written report for Globox

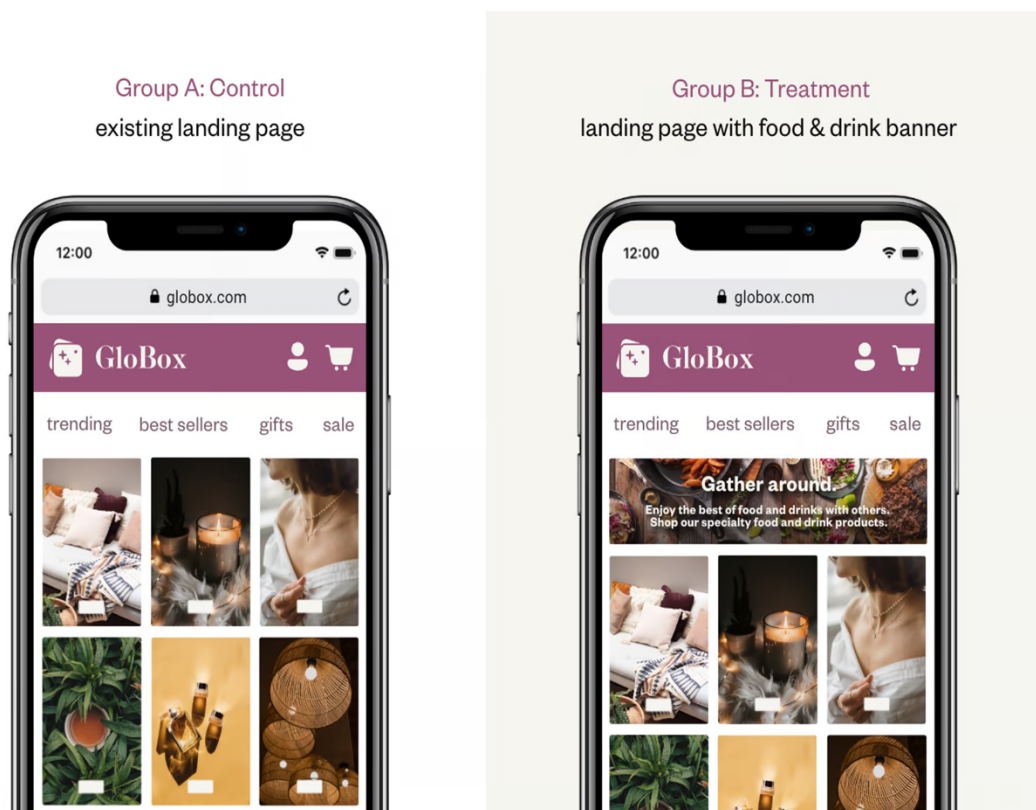


Written by: Bijal Patel
Date: September 12th, 2023

Summary:

GloBox is an online marketplace that specializes in sourcing unique and high-quality products from around the world and is primarily known amongst its customer base for boutique fashion items and high-end decor products. However, their food and drink offerings have grown tremendously in the last few months, and the company wants to bring awareness to this product category to increase revenue.

The Company's Growth team decides to run an A/B test that highlights key products in the food and drink category as a banner at the top of the website. The control group does not see the banner, and the test group sees it as shown below:



The setup of the A/B test is as follows:

1. The experiment is only being run on the mobile website.
2. A user visits the GloBox main page and is randomly assigned to either the control or test group. This is the join date for the user.
3. The page loads the banner if the user is assigned to the test group and does not load the banner if the user is assigned to the control group.

4. The user subsequently may or may not purchase products from the website. It could be on the same day they join the experiment, or days later. If they do make one or more purchases, this is considered a “conversion”.

Task:

Our task is to analyze the results of the A/B test and provide a recommendation to the stakeholders about whether GloBox should launch the experience to all users.

Experiment’s Sample Size:

- Total Users: 48,943
- Control Group: 24,343
- Treatment Group: 24,600

Goal of the Experiment:

- Should we Launch the Banner?
- Do we have enough information to Decide?
- Should we iterate on this experience and test it again?

The project features three stages

1. Extracting the user-level aggregated dataset using SQL.
2. Analyzing the A/B test results using statistical methods in spreadsheets
3. Visualizations in Tableau.

Context:

1. Extracting the user-level aggregated dataset using SQL:

This initial stage involved extracting the user-level aggregated dataset using SQL in Beekeeper Studio. I connected the GloBox database to Beekeeper Studio and crafted the SQL query outlined below. Successfully obtaining the required dataset marked a crucial step in advancing the experiment.

Below is the SQL query:

```
SELECT u.*, g.device AS users_device, g.group AS test_group, CASE WHEN a.spent > 0 THEN 1 ELSE 0  
END AS Converted, COALESCE(SUM(a.spent), 0) AS Total_spent  
FROM users u
```

LEFT JOIN groups g ON u.id = g.uid

LEFT JOIN activity a ON u.id = a.uid

GROUP BY u.id, u.country, u.gender, g.device, g.group, converted ;

2. Analyzing the A/B test results using statistical methods in spreadsheets:

In the second stage, I imported the downloaded CSV dataset file into Google Sheets and performed necessary inferential statistics using spreadsheets.

- Conversion rate analysis between two test groups: From Statistical Analysis of Conversion rate for Both Test Group, **p Value is $p = 0.0001114119853$ which is <0.05 , Which means it Rejects Null Hypothesis**, meaning there is a significance difference in conversion rate of both test group.

- 95% confidence interval for conversion rate between two groups: Difference in conversion rate was found to be between 0.0034% and 0.0106%.

Statistical Analysis of Conversion rate For Test Groups

Descriptive statistics	Notation	Test Group A	Test Group B
Sample Size	n	24,343.00	24,600.00
Converted Users	x	955.00	1,139.00
Sample Proportion of Success/Conversion Rate	p^{\wedge}	0.04	0.05
Standard Error	SE	0.00	0.00
Pooled Proportion	p^{-}	0.04	
Conversion Rate Difference	%/%	0.01	
Total Standard Error	(SE)	0.00	
Degree of Freedom	$(n1-1)+(n2-1)$	48,941.00	
Z Score	z	3.86	
Critical Value	z^{*}	1.96	
Q1- P-value	p	0.00	
Level of Significance	α	0.05	

Q2- Confidence Interval	CI	0.00	0.01
-------------------------	----	------	------

- Average amount spent per user between the two groups: From Statistical Analysis of Average amount spent for Each test group, **P value is 0.9438557529, which is >0.05 , that indicated it Fail to Reject Null Hypothesis**, which means there is no significance difference in the average amount spent in both Test Groups.

- 95% confidence interval for the difference in the average amount spent between two groups: Difference in average amount spent was found to be between -0.438 and 0.471.

Statistical Analysis of Average Amount Spent for Test Group

Descriptive statistics	Notation	Group A	Group B
Sample Size	n	24,343.00	24,600.00
Average Amount Spent	\bar{x}	3.37	3.39
Standard Deviation	s	25.94	25.41
Total Average Spent	\bar{x}	3.38	
Pooled Standard Deviation		0.05	
Standard Error Total	(SE)	0.23	
Degree of Freedom	(n1-1)+(n2-1)	48,941.00	
Mean Difference	$\bar{x}_1 - \bar{x}_2$	0.02	
Critical Value	z^*	1.96	
Confidence level	$1 - \alpha$	0.95	
T-value	T	0.07	
Q3- P-value	P	0.95	
Level of Significance	α	0.05	

Q4- Confidence Interval	CI	-0.44	0.47
-------------------------	----	-------	------

Summary of other statistical analysis for conversion rate:

- Sample size for Control: 24343
- Sample size for Treatment: 24600
- Converted users for Control :955
- Converted users for Treatment: 1139
- Conversion rate for Control: 0.0392
- Conversion rate for Treatment: 0.0463
- Standard Error for Control: 0.001244
- Standard Error for Treatment: 0.001339

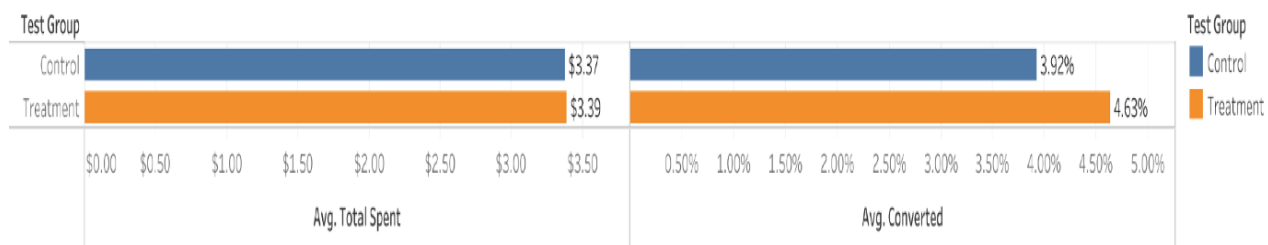
- Pooled **Proportion:** 0.0427
- Degree of freedom: 48941
- Z score: 3.864
- Critical Value: 1.960

Summary of other statistical analysis for average amount spent:

- Sample size for Control: 24343
- Sample size for Treatment: 24600
- Average amount spent for Control: 3.374518468
- Average amount spent for Treatment: 3.390866946
- Pooled Standard Deviation: 0.04580993
- Total Standard Error: 0.2321405588
- Degree of Freedom: 48941
- Mean Difference: 0.01634847796
- Critical Value: 1.960012458
- Confidence Level: 0.95
- T Value: 0.07042491
- P Value: 0.9438557529

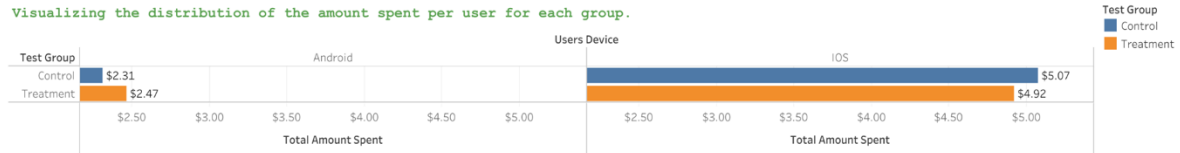
3. Visualizations in Tableau:

visualization to compare the conversion rate and average amount spent between the test groups.



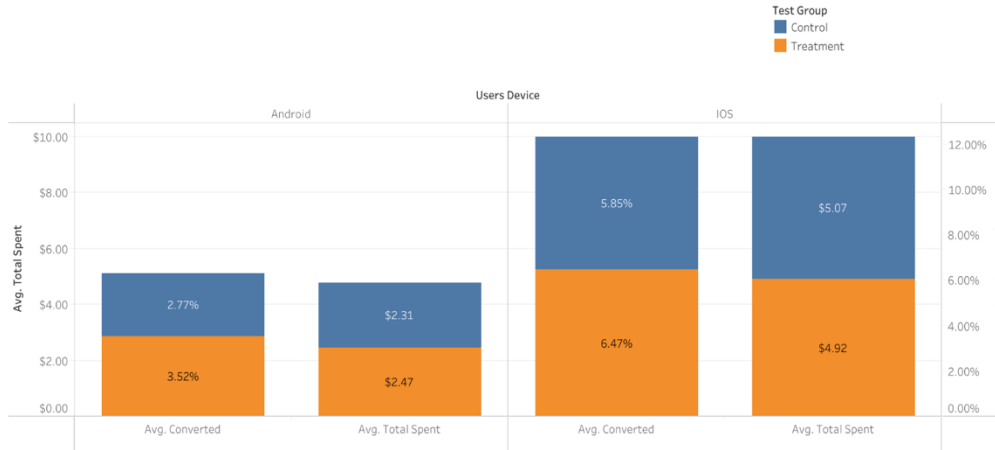
- Conversion rate: This visualization shows us that more users are converted in the treatment group with conversion rate 4.63% as compared to the control group with conversion rate 3.92%.
- Average total spent: However, not a big difference in Average amount spent was observed between the treatment group with \$3.39 average spent and control group with \$3.37 average spent.

Visualizing the distribution of the amount spent per user for each group.



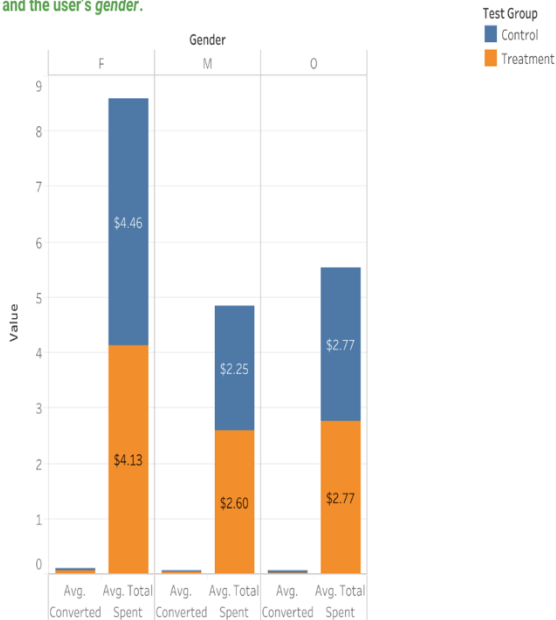
- **iOS :** This visualization shows us that iOS device users spent more on an average from the control group with average spent of \$5.07 as compared to treatment group with average spent of \$4.92.
- **Android:** In contrast to iOS, Android device users spent more on an average from Treatment group with average spent of \$2.47 as compared to control group with average spent of \$2.31.

visualizations to explore the relationship between the test metrics (conversion rate and average amount spent) and the user's device .



- This visualization shows us that iOS device users have higher conversion rate and average amount spent for both the test groups as compared to Android device users.

visualizations to explore the relationship between the test metrics (conversion rate and average amount spent) and the user's gender.



- This visualization shows us that Females have higher average amount spent as compared to all other gender across both the test groups. Control group: \$4.46 and Treatment group: \$4.13.

Recommendation:

From Statistical Analysis as well as Graphical visualization it is very clear that there is a significant increase in Conversion rate in the treatment group that see's the banner on their device. However, no significant increase was observed in the average amount spent between both the test groups.

Since we did see a significant increase in Conversion rate but do not have enough evidence to prove the increase in average amount spent between the two groups. At this point, instead of coming to a decision of launching the new banner or not, I recommend iterating the experiment with a larger sample size or for an extended period of time to get to a proper decision of Launching the banner or not.

Appendix

1. Extracting the user-level aggregated dataset using SQL.

```
SELECT u.*, g.device AS users_device, g.group AS test_group, CASE WHEN a.spent > 0 THEN 1 ELSE 0 END
AS Converted, COALESCE(SUM(a.spent), 0) AS Total_spent
FROM users u
```

```
LEFT JOIN groups g ON u.id = g.uid
LEFT JOIN activity a ON u.id = a.uid
GROUP BY u.id, u.country, u.gender, g.device, g.group, converted ;
```


Analyzing the A/B test results using statistical methods in spreadsheets:

2. Below is the link to the spreadsheet that contains all the statistical calculation for the experiment.

<https://docs.google.com/spreadsheets/d/1JAFeQUyIYPWy6DXzJiPgcTrLzatWvbZwhqdWuIV5CB8/edit?usp=sharing>

3. Visualizations in Tableau.

- Below is the link to Tableau Dashboard:

https://public.tableau.com/views/GloboxVizualizationinTableau/Dashboard5?:language=en-GB&:display_count=n&:origin=viz_share_link

- Below is the link to the Visual Presentation slides in Google slides with explanation.

Globox Project

https://docs.google.com/presentation/d/1WVEBUi_iI5sJZfBC4k3VEziZzNn9VF4Ja_ZpWuHATiI/edit?usp=sharing