

# A/B test GloBow – Banner

## Introduction

This written report explains the steps that were taken to be able to draw a conclusion regarding whether adding a banner at the top of an e-commerce web shop will increase revenue. Additionally, the corresponding results and conclusions will be listed.

## Experiment Information

GloBox is an online web shop known for its boutique fashion items and high-end décor. Since their food and drink offerings have grown tremendously in the last few months, GloBox wants to bring awareness to this category and hopefully increase revenue.

To test if the changes made will yield the desired results, an A/B test is run.

Every user who uses the mobile website during the experiment will be assigned to either group A or Group B.

When a user in group A visits the website, he will see no changes.

For users in group B a banner at the top of the website is shown which promotes the Food and Drink category.

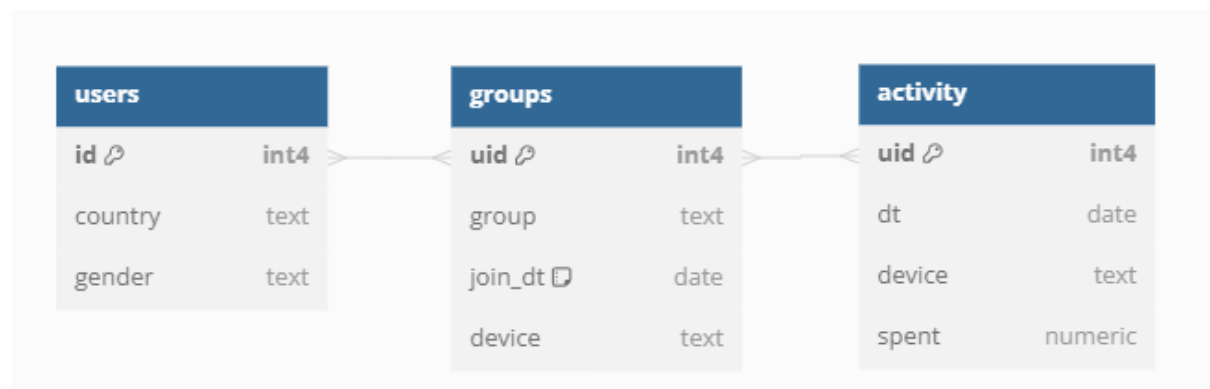
The experiment will be run for 13 days starting from 25-01-2023 until 06-02-2023

## Gathering data

The data will be gathered with the help of the SQL database and will be exported to a csv.

Only data within the time period of the experiment will be exported.

The Data is spread over the following tables:



The following columns will be needed:

1. Uid – user\_id
2. Country
3. Gender
4. Device – Android or Ios (A/I)
5. Group – The experiment group users are in (A: control, no changes - B: experiment, including banner)
6. Converted – Every user who bought at least one product is converted (Y: yes, N: not)

## 7. Total\_spent – The total amount spent per user in the given time period

SQL query to obtain data:

```
SELECT DISTINCT uid, country, gender, device, sub.group, converted, total_spent
FROM(
SELECT g.uid, u.country, u.gender, g.device, g.group,
CASE WHEN spent > 0 THEN 'Y'
ELSE 'N' END converted,
SUM(spent) OVER(Partition BY g.uid) total_spent
FROM activity a
FULL JOIN groups g USING(uid)
LEFT JOIN users u ON g.uid=u.id) sub
```

Note: The distinct statement is used to drop any duplicates. Beforehand was checked if duplicated ids contained different values by the following query. This was not the case.

```
WITH more_than_once AS(
SELECT g.uid, COUNT(*) OVER(PARTITION BY uid) uid_count
FROM activity a
FULL JOIN groups g USING(uid)
LEFT JOIN users u ON g.uid=u.id
)

SELECT uid, country, gender, device, sub.group, converted, COUNT(*)
FROM(
SELECT g.uid, u.country, u.gender, g.device, g.group,
CASE WHEN spent > 0 THEN 'Y'
ELSE 'N' END converted,
SUM(spent) OVER(Partition BY g.uid) total_spent
FROM activity a
FULL JOIN groups g USING(uid)
LEFT JOIN users u ON g.uid=u.id) sub
WHERE uid IN(SELECT uid FROM more_than_once WHERE uid_count > 1)
GROUP BY 1,2,3,4,5,6
HAVING COUNT(*) < 2
ORDER BY uid
```

## Significance testing

### Overview

For this project two significant test have been carried out to answer the following questions:

1. Does adding the banner increase the conversion rate?
2. Does adding the banner increase revenue?

The method used as well as the result will be explained here. However, the test themselves are carried out inside a Jupyter Notebook with the help of Python and packages as Pandas, Numpy, and Matplotlib. For details look in the Jupyter Notebook file named: GloBox A-B test – banner.ipynb

## 1. Does adding the banner increase the conversion rate?

H<sub>0</sub>: Adding banner with key products does not increase the conversion rate

H<sub>a</sub>: Adding banner with key products does increase the conversion rate

H<sub>0</sub>: P<sub>exp</sub> - P<sub>con</sub> ≤ 0

H<sub>a</sub>: P<sub>exp</sub> - P<sub>con</sub> > 0

α = 0.05

Z-test

$$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

This resulted in a z-value of 5.712 which is far above the one tailed critical z-value of 1.65 given α=0.05. Thereby the null hypothesis is rejected and we accept the alternative hypothesis.

With a z-value of 5.712 we are more than 99.99% confidence that adding a banner increases the conversion rate.

## P-value per country

The mean difference per country is calculated with the help of the groupby function. After obtaining the difference a dictionary is created where every country is added as a key and the corresponding z-value as value using the formula listed above. With the help of scipy.stats.norm.cdf function the z-value is transformed to a percentage which is printed below with its corresponding country ID.

Some countries do not have a significance level above 95%. That means that in those countries we can't reject the null hypothesis with confidence when only looking at the country specific data.

## Bootstrap

Also, a bootstrap is run to simulate the distribution of means and obtain a p-value. For 10.000 times a random sample is drawn with replacement in size of the provided data. Then The difference in mean of both groups is calculated and added to a list.

Now a list, called "null\_vals", is created where with the help of np.random.normal function values are added expected by a normal distribution given the size and standard deviation. When comparing both arrays, diff array and "null\_vals" array, we can obtain a p-value by calculating the mean of the given Boolean list. In this case 0.0044 which is below our alpha.

Given this result in we can conclude once again that we are confident to reject the Null Hypothesis.

## 2. Does adding the banner increase revenue?

H<sub>0</sub>: Adding banner with key products does not increase mean revenue per customer

H<sub>a</sub>: Adding banner with key products does increase mean revenue per customer

H<sub>0</sub>: u<sub>exp</sub> - u<sub>con</sub> ≤ 0

H<sub>a</sub>: u<sub>exp</sub> - u<sub>con</sub> > 0

α = 0.05

## Bootstrap

Again, a bootstrap is run. Only this time it compares the differences of the average amount spent per user over the two groups.

For 1000 times a sample is drawn of the data with replacement and the same size as the original data. After this the means for either group is calculated and the difference between those means added to the list *diffs*. The list is plotted to make sure data is normal distributed. Since the null hypotheses expects to have a normal distribution around 0, a list is created with a simulated distribution of this kind. After comparing the “diff list” with the created “null-valls” list a p-value of 0.5 is obtained showing that the distribution overlap rather frequently.

We can't reject the null hypothesis.

## Z-test

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2 / n_X + \sigma_Y^2 / n_Y}} \sim \mathcal{N}(0, 1)$$

This resulted in a z value of 0.070 which does not surpass the required z-value of 1.68 to reject the null hypothesis.

We fail to reject the Null Hypothesis. Including the banner promoting the foods and drinks category does not significantly increase revenue.

## P-Value per country

Per country and group the means and standard deviations were obtained and the resulting z-value was added to a list. The z-value was then recalculated to be shown as significance levels. Besides Great-Britain there was no country which showed a significant increase in revenue.

## Conclusion

The conversion could be significantly increased by adding the banner. However, this implementation did not significantly increase our revenue. Also, what needs to be mentioned is, that it didn't decrease the revenue either. Since increasing revenue was the goal of this banner, I do not advise introducing it globally.

Looking at Great-Britain merely, there was a significant increase in both revenue as conversion. Running further tests to confirm results would be advisable.

## Links

[Tableau Public](https://public.tableau.com/app/profile/torben.mulitze/viz/GlowboxAB/Dashboard1) <https://public.tableau.com/app/profile/torben.mulitze/viz/GlowboxAB/Dashboard1>