

GloBox new homepage Experiment Report

08/20/2023

By : Jugurtha Hachemi

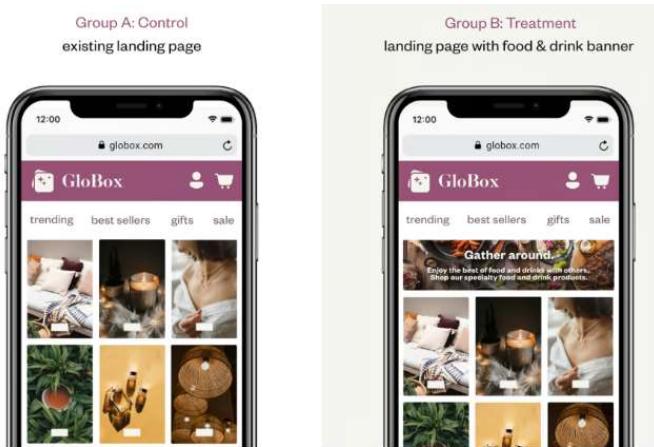
Summary

Globox company has created a new homepage that wants to launch with new features, we as a data analysts team, we extracted the user-level aggregated dataset using SQL. And analyze the A/B test results using statistical methods in spreadsheets, where we run a hypothesis test for the conversion rate and for the mean. We also did a visualizations in Tableau.

Context

We ran an A/B test with a new design for the homepage to see if it will increase revenue. You can see the difference between the two designs below. The Control group as the existing design,

And the Treatment group as the new design.



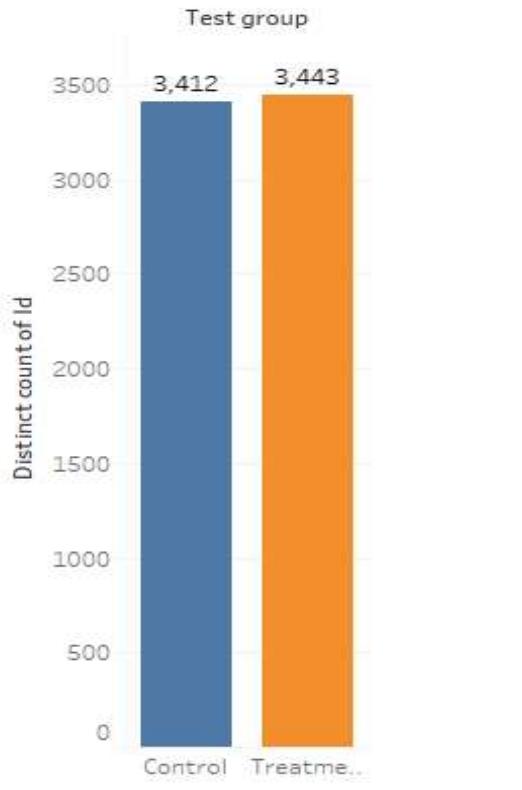
- The experiment ran between 01/25/2023 and 02/06/2023. There were 48943 total users, 24343 users in the control group and 24600 users in the treatment group.

users by group and device

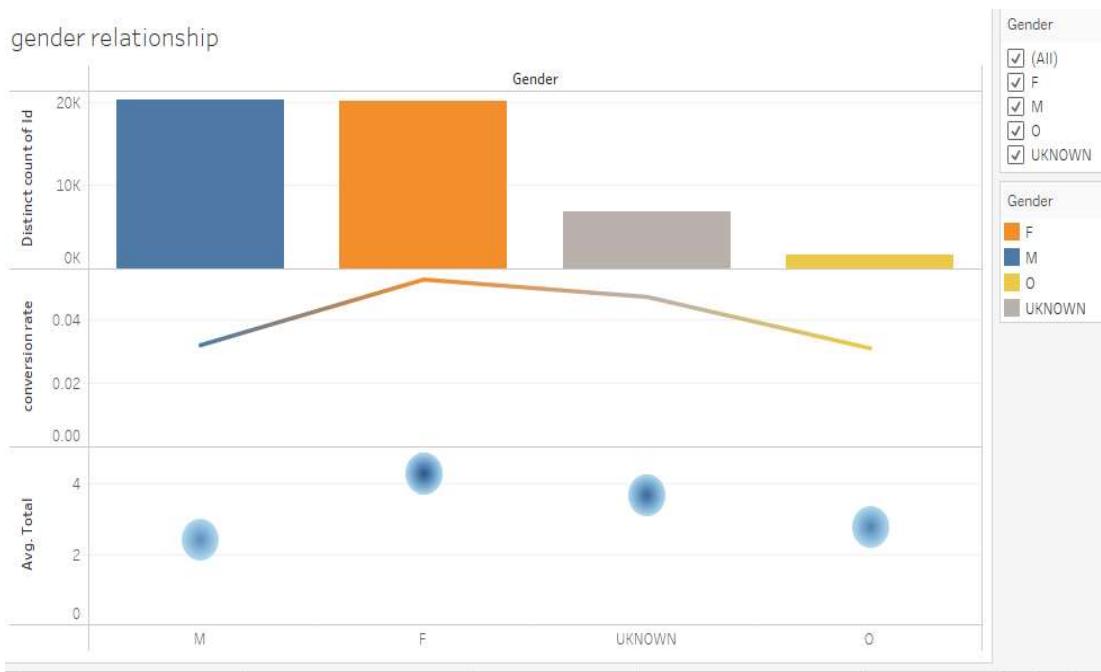
Test group		Device			Grand Tot..
		ANDROID	IOS	no info	
Treatment	Distinct count of Id	15,235	9,218	147	24,600
	user converted	537	596	6	1,139

- The histogram below shows the missing gender in each group

missing gender by group

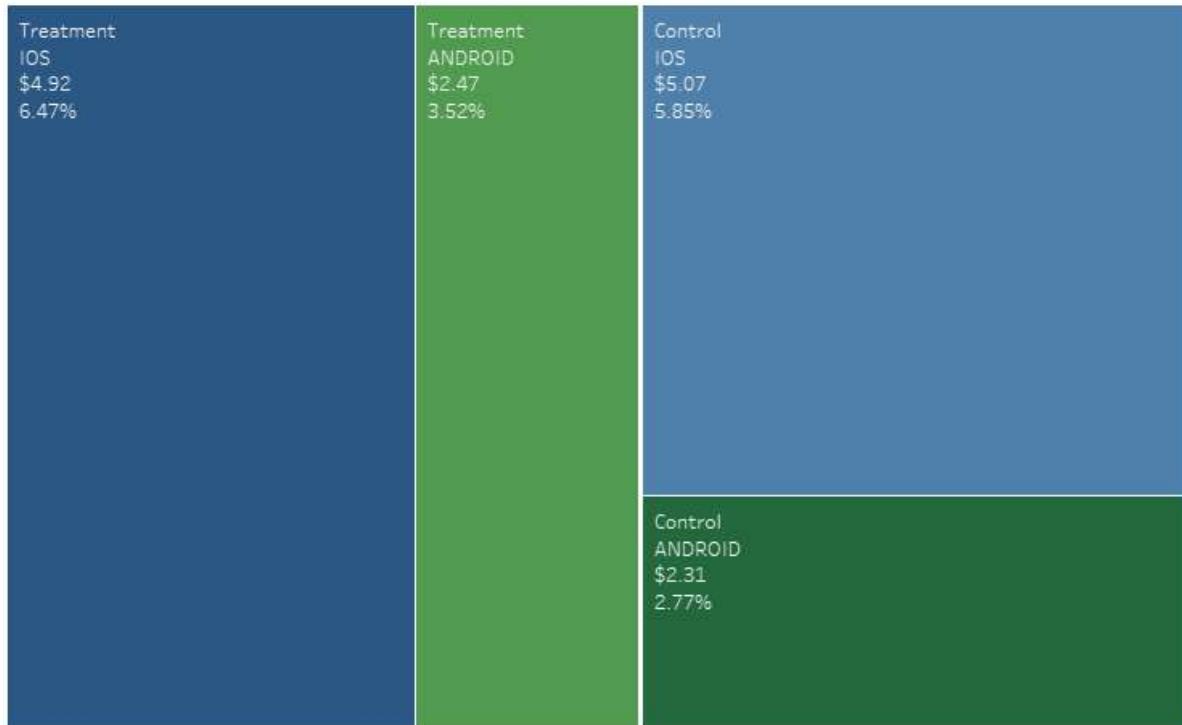


- Below the visualizations that shows the gender relation based on conversion rate and AVG total spent



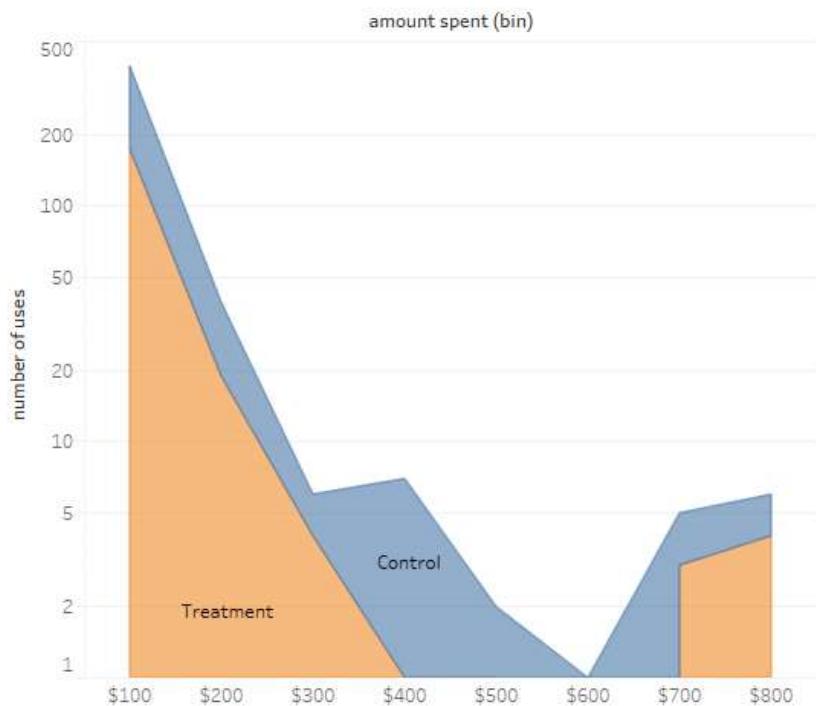
- Another visualization that shows the relation between conversion rate and AVG amount spent and devices(ANDROID & IOS)

relationship between cRT and AVG and user device



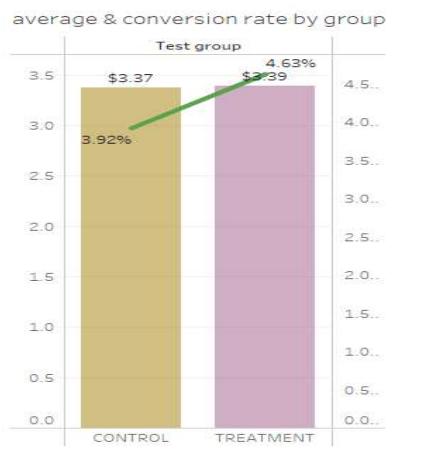
- This is the distribution of amount spent by user in each group

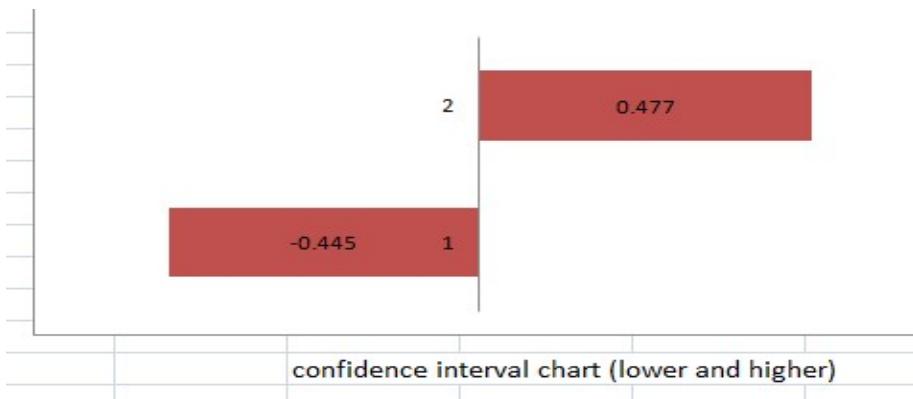
distribution of amount spent by user



- **RESULTS**

In order to determine if there is a difference in the average amount spent by user in each group, we ran a hypothesis test of the mean. We did not see a statistically significant difference between the two groups at the 5% significant level where ($p=0.94$). The 95% confidence interval for the difference in the average amount spent per user between the two groups is (-0.439, 0.471).





- To determine if the conversion rate between the two groups (control and treatment) is equal, we ran a hypothesis test of the proportion. We did see there was a statistically significant difference between the two groups at 5% significant level where ($p=0.0001$). The 95% confidence interval for the difference in the conversion rate between the two groups is (0.0034, 0.0106).

Confidenc intervl chart



• RECOMMENDATION

By conducting our hypothesis tests (proportion, conversion rate) and analyzing our results we recommend you to not launch the new home page because we did not see that it will increase the revenue

Below are the codes with explanation that we use on our experiment

We started by extract our data and cleaning it, using SQL commands

- We started by checking if a user can show up more than once in the **activity** table by doing multiple purchases

2233

```
select count(uid)- select count(Distinct(uid))
```

```
From activity;
```

the result was 139 users

- We joined the **users** table to the **activity** table by using left join

```
select *
```

```
From users
```

```
Left join activity
```

```
ON users.id = activity.uid;
```

- We used the COALESCE function to fill in NULL values

```
COALESCE(gender,'UNKNOWN') as gender
```

- We used MAX(), and MIN() function to get the start and end dates of the experiment (01/25/2023 end 02/06/2023)

```
SELECT MAX(dt), MIN(dt)  
from activity;
```

- By using count(distinct) function we got the total users in the experiment (48943)

```
Select count(DISTINCT uid)
```

```
From groups;
```

- we get the number of users in each groups (control (24343), treatment(24600)

Control

```
select gr.group,count(*)  
FROM groups gr  
WHERE gr.group ='A'  
GROUP BY gr.group;
```

Treatment

```
select gr.group,count(*)  
FROM groups gr  
WHERE gr.group ='B'  
GROUP BY gr.group;
```

- We calculated the conversion rate for all users (4.28%)

```
select cast(count(distinct uid) as float)/ cast( count(distinct id) as float) total_conversion_rate  
from activity act
```

```
left join users us
```

```
on act.uid=us.id ;
```

- we calculated the conversion rate for each group (control 3.92%, treatment 4.63%)

```
=>control
```

```
select cast(count(distinct act.uid) as float)/ cast( count(distinct gr.uid) as float)
```

```
conversion_rate_control
```

```

from activity act
left join groups gr
on act.uid=gr.uid
WHERE gr.group='A';
=> treatment
select cast(count(distinct act.uid) as float)/ cast( count(distinct gr.uid) as float)
total_conversion_rat_control
from activity act
left join groups gr
on act.uid=gr.uid WHERE gr.group='B'

```

- we looked for the average amount spent per user in control and treatment (\$3.373, \$3.391)
WITH my_table as
(select

```

DISTINCT(id),country,COALESCE(gender,'UNKNOWN'),COALESCE(gr.device,'UNKNOWN'),gr.group as conv ,
SUM(COALESCE(spent,0.0)) as total
from users
left join groups gr
on users.id=gr.uid
left join activity act
on act.uid=gr.uid
GROUP BY id,country,gender,act.device,gr.group)

```

```

select round(avg(total),3)
from my_table
where my_table.conv = 'A'; / where my_table.conv = 'B';

```

- We got all the data we need for the next phase of testing (user ID, the user's country, the user's gender, the user's device type, the user's test group, whether or not they converted (spent > \$0), and how much they spent in total (\$0+)).

```

select DISTINCT(id),country,COALESCE(gender,'UNKNOWN') as gender,COALESCE(gr.device,'no info')
as device, gr.group as group_test,SUM(COALESCE(spent,0.00)) as total

from groups gr
left join activity act
on gr.uid=act.uid
left join users on users.id=gr.uid
GROUP BY users.id,country,gender,gr.device,gr.group

```

- Tableau link to visualization
<https://public.tableau.com/app/profile/jugurtha.hachemi/viz/jugurthagloboboxexperimant/Story1?publish=yes>
- Tableau link to CONVIDENCE INTERVAL & NOVELTY EFFECT
<https://public.tableau.com/app/profile/jugurtha.hachemi/viz/confidenceintervalnoveltyeffectintablau/CONFIDENCintvNovltefct?publish=yes>

