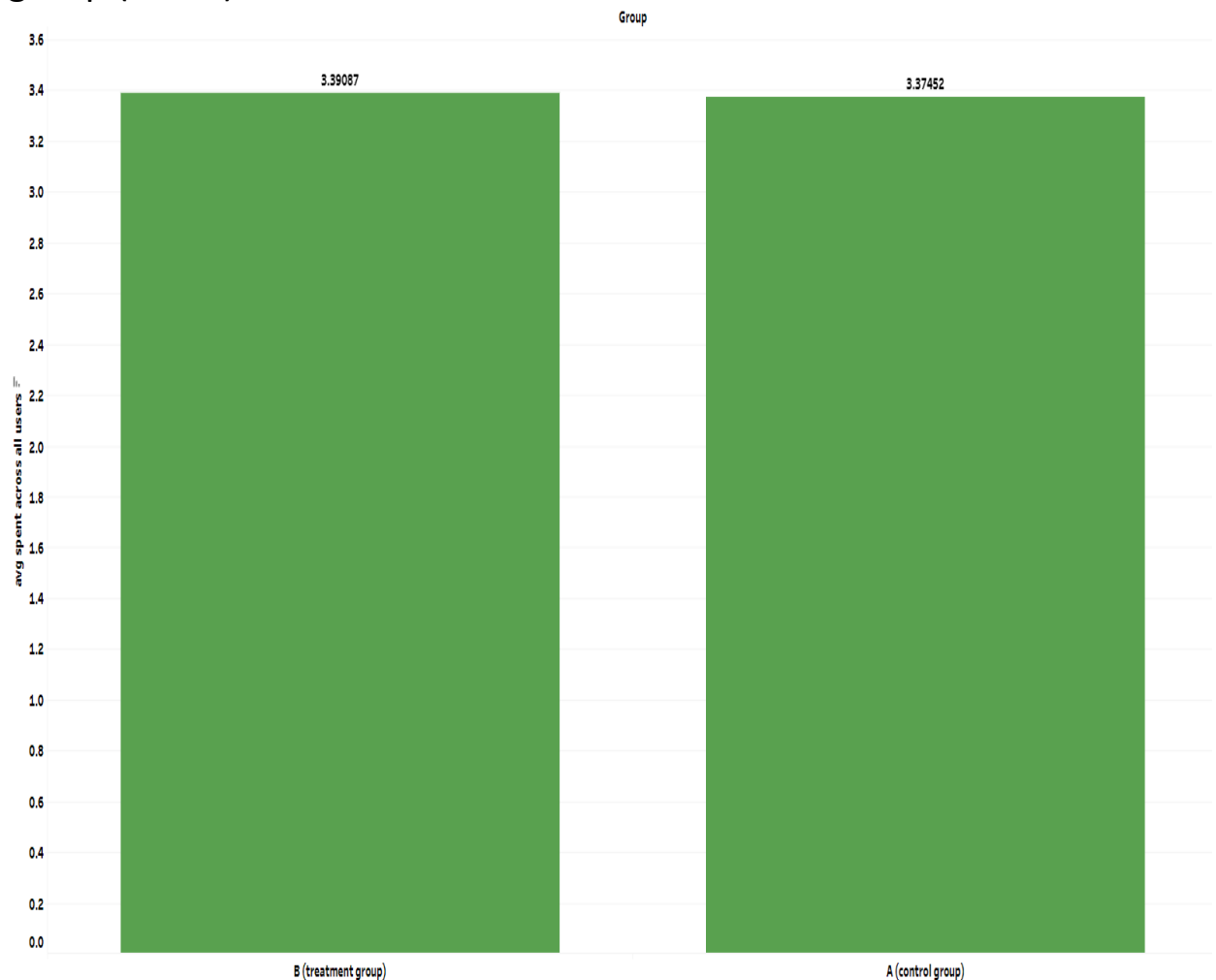# GLOBOX A/B TEST REPORT

Avg spent per customer point of view :

Confidence intervals

To get our estimate TRUE average spent per customer in our two groups first of all we need to cauculate the average amount spent per user in the treatment (3.390$) and control group (3.374).

After that I set the confidence level to 95 which means that after the calculations we can be 95% confident of our results which is that the true avg spent per customer is :

Control : (3.049,3.700) in other words, we can be 95% confident that the true revenue per customer in group A is between 3.049 and 3.700 dollars.

Treatment : (3.073,3.708) in other words, we can be 95% confident that the true revenue per customer in group B is between 3.073 and 3.708 dollars.
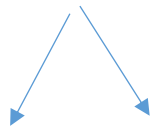
CAUCULATIONS:

CONTROL :

standard error : 25.936 / sqrt(24343) = 0.170

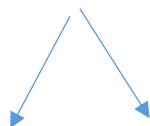margin of error : 1.960 * 0.170 = 0.332

CI: 3.374 +- 0.332

(3.049   3.69)

Treatment :

standard error : 25.414 / sqrt(24600) = 0.162

margin of error : 1.960 * 0.170 = 0.317

CI: 3.391 +- 0.317

(3.073   3.708)

Hypothesis test

After calculating the above we need to check if the change in the GLOBOX website is worth launching based on the raw revenue. Based on looking at the numbers alone we can see that the difference between the two groups is insignificant. And to prove that we would perform a hypothesis test given our alpha is 0.5 which is the 100% - confidence level (95%).

After setting our significance level (alpha) we conclude that we should perform a Two-sample t-test. After that we define our null and alternative hypothesis:

H0 : mean 1 = mean 2 .

H1 : mean1 not equal to mean 2 .

After that we cauculate that our p value is 0.944 and we compare it to our alpha :   0.944 > 0.5 which tells us that our difference is statistically insignificant and that we fail to reject our null hypothesis. In other words that based on the revenue per customer we should not launch the change in the website .
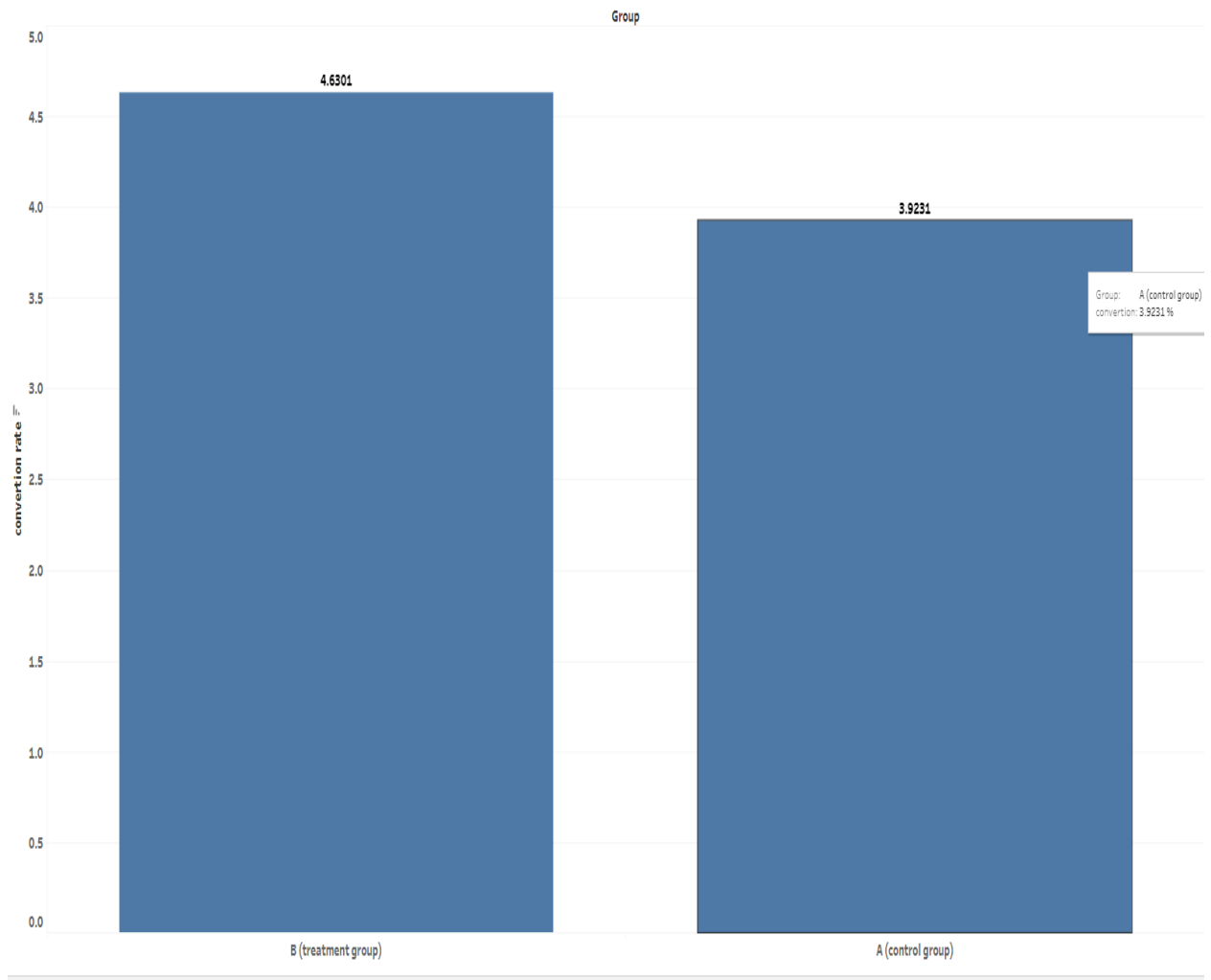
Calculations :

T test statistic : 3.390 – 3.374 / sqrt ((25.414/24600) + (25.936/24343)) = 0.045

P value : 0.944

Conversion spent per customer point of view :

First of all I calculated the conversion rate for the control (3.92%) and for the treatment (4.63%)

After that I set the confidence level to 95 which means that after the calculations we can be 95% confident of our results which is that the true conversion rate in each group is :

Control : (0.0368 , 0.0417) in other words, we can be 95% confident that the conversion rate in group A is between 3.68% and 4.17%.

Treatment : (0.0437 , 0.0489) in other words, we can be 95% confident that the conversion rate in group B is between 4.37% and 4.89% .

Hypothesis test

After calculating the above we need to check if the change in the GLOBOX website is worth launching based on our conversion rate. Based on looking at the numbers alone we can see that the difference between the two groups is 1% which might not seem allot but it is significant. And to prove that we would perform a hypothesis test given our alpha is 0.5.

After setting our significance level (alpha) we conclude that we should perform a Two-sample z-test. After that we define our null and alternative hypothesis:

H0 : both proportions are equal .

H1 : the proportions differ from each other .

After that we cauculate that our p value is 0.0001 using the Two-sample z-test with pooled proportion formula and we compare it to our alpha :   0.001 < 0.5 which tells us that our difference is statistically significant and that we reject our null hypothesis. In other words that based on the conversion rate we should launch the change in the website .

- I didn't show the calculations in the conversion part because of the report length but I used the Two-sample z-interval with unpooled proportions to cauculate the CI and p value .

Conclusion

In my opinion we should launch the change for those reasons :

1. We can see that the treatment group exceeds in both aspects of revenue and conversion although we cant be sure about the revenue we can look only on the conversion rate which has a big enough difference.
2. In my point of view conversion rate is more important than raw revenue because conversion means that there is a big chance that we have another happy customer that will probably use the service of GLOBOX company in the future and as we convert more that's an investment to the future .
3. The change in the website is affordable and does not cost a lot compared to other changes .

Codes and links

Avg spent:

WITH total_per_user AS (

SELECT uid, SUM(spent) AS total_spent

FROM activity

GROUP BY uid

)

SELECT g."group", AVG(COALESCE(total_spent, 0))

FROM groups g

LEFT JOIN total_per_user a ON g.uid = a.uid

GROUP BY g."group";

Standard deviation per group:

```
WITH total_per_user AS (

SELECT uid, SUM(spent) AS total_spent

FROM activity

GROUP BY uid

)

SELECT g."group", stddev(COALESCE(total_spent, 0))

FROM groups g

LEFT JOIN total_per_user a ON g.uid = a.uid

GROUP BY g."group";
```

Conversion rate :

```sql
WITH t1 AS

(

SELECT COUNT (id) as num_a, g.group AS ga

FROM users u

LEFT JOIN groups g

ON g.uid = u.id

WHERE g.group = 'A'

GROUP BY 2


)
,t2 AS

(

SELECT COUNT (id) as num_b, g.group AS gb

FROM users u

LEFT JOIN groups g

ON g.uid = u.id

WHERE g.group = 'B'

GROUP BY 2
```

```
)

, t3 AS

(

SELECT COUNT (DISTINCT a.uid) AS aa
,g.group

FROM groups g

LEFT JOIN activity a

ON a.uid = g.uid

WHERE g.group = 'A'

GROUP BY 2


)


,t4 AS

(

SELECT COUNT (DISTINCT a.uid) AS bb
,g.group

FROM groups g

LEFT JOIN activity a
```

ON a.uid = g.uid

WHERE g.group = 'B'

GROUP BY 2

)

SELECT   aa/CAST(num_a AS float) * 100 AS convertion_A , bb/CAST(num_b AS float) * 100 AS convertion_B

FROM t1,t2,t3,t4

Tableau charts :

https://public.tableau.com/app/profile/rwad7818/viz/ABtest_16838101886460/Story1