

# **Written Report**

## **GloBox A/B Testing Analysis**

### **Objective: Experimenting a New Food & Drink Landing Page Banner**

**Analyst : Djamel Fitas**

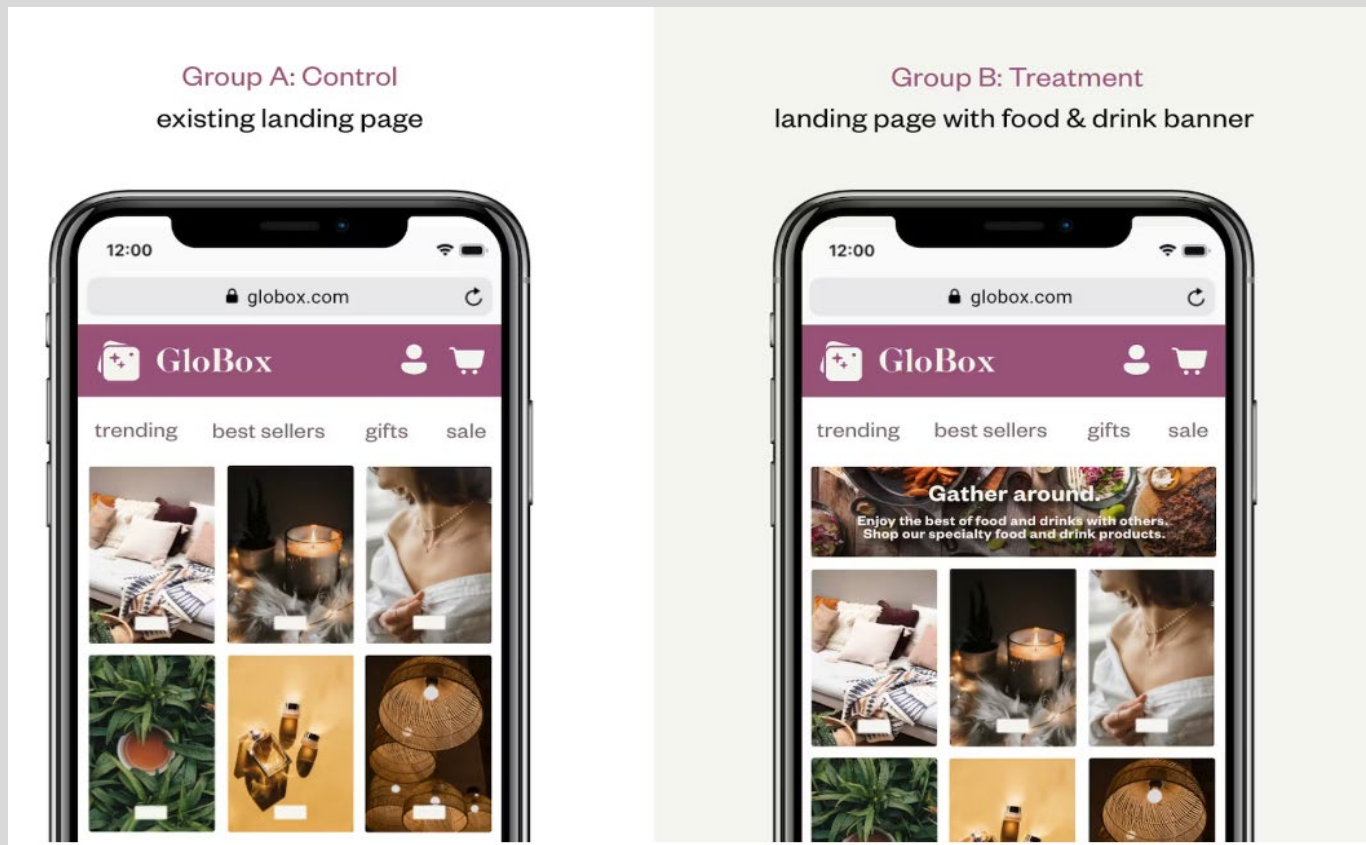
**TL, DR**

#### **Recommendation**

**The new app version increased conversions by 18%, hence the new updated app interface should be released to the public.**

# Summary

A/B testing was carried out for a landing page banner displaying food and product category. In the trial, users were placed into two groups: A and B, which acted as the control and treatment, respectively. Tracking the proper data for each group, such as conversion rates and spending amount, to evaluate how effective the banner is during a specific time period.



## Experiment Context

We intended to do A/B testing and make business decisions based on the results in order to determine if the food and beverages banner would be featured on the app. Assigning users equally to control and treatment groups and tracking their purchasing activities with the option to convert after seeing the banner.

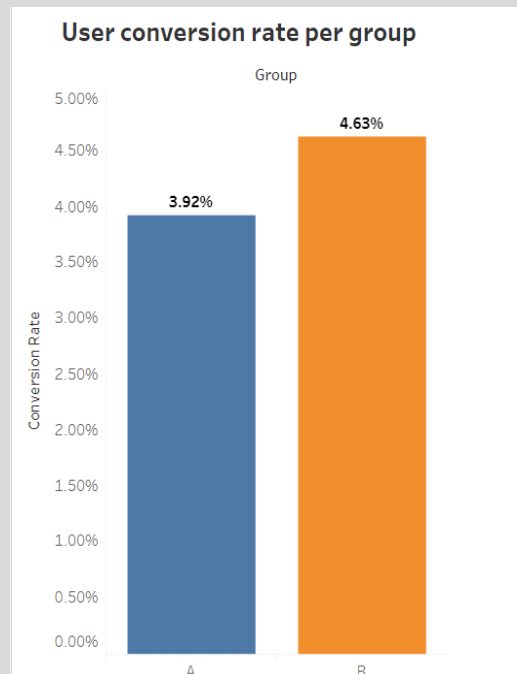
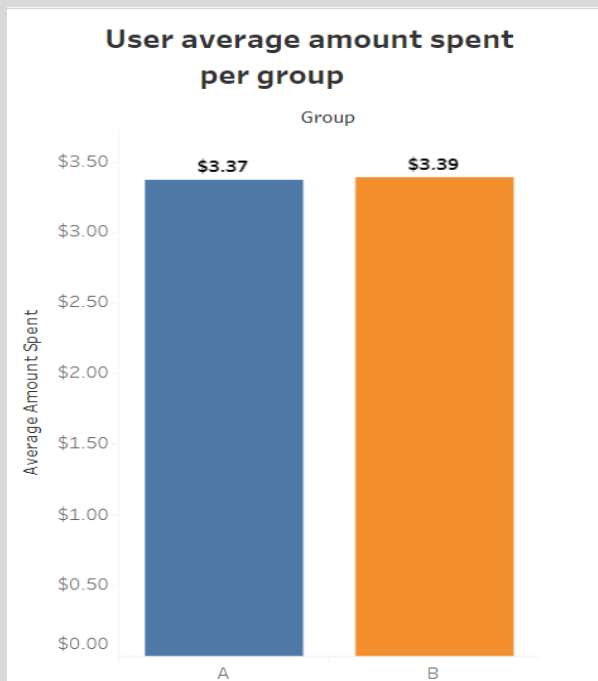
The conversion rate, defined as the percentage of users who made a purchase during the experiment, and the amount spent per user were deemed critical metrics and recorded for each individual between January 25th and February 6th, with a total number of users reaching 49k divided 50/50 to each group.

# Hypothesis testing Results

The sample user analysis revealed an 18% increase in conversions and a 0.6% increase in total amount spent, indicating that the banner did not influence as many users to convert as anticipated.

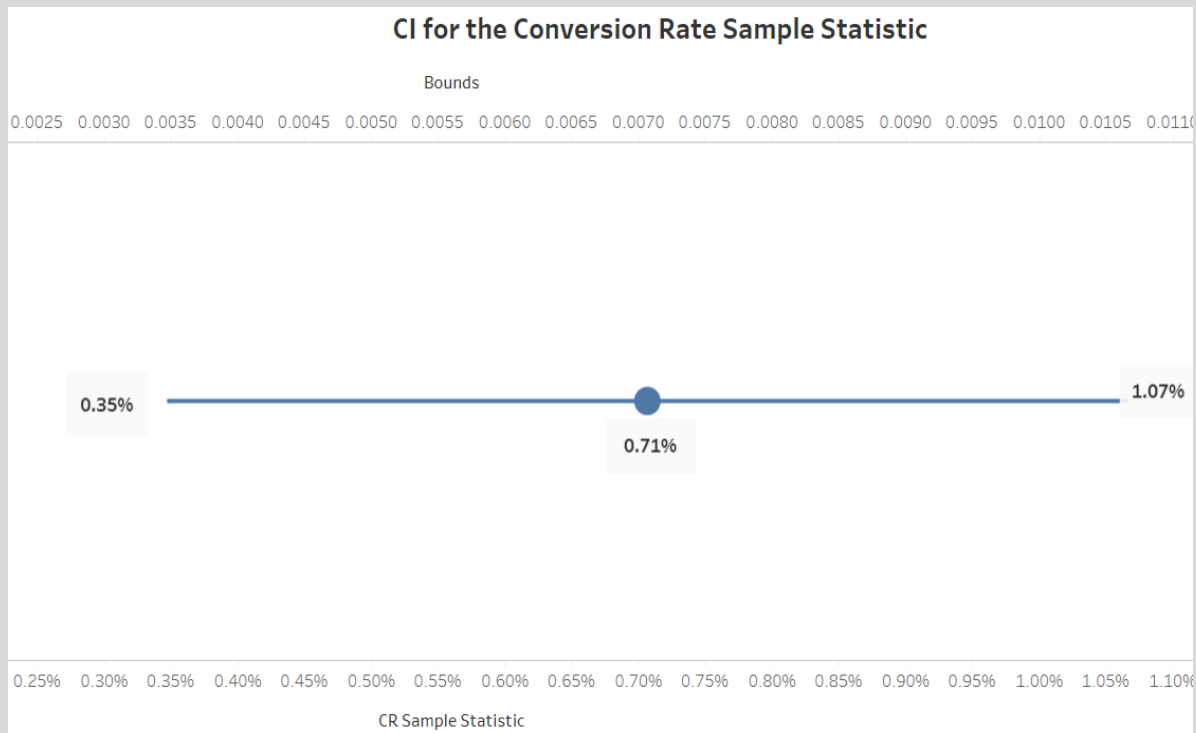
Based on our sample size and the results of the hypothesis testing performed on the data we gathered, we found that the conversion rate for Group A was 3.92% and for Group B it was 4.63% with an 18% lift. Statistically, this result is very significant, with a p-value of 0.0001. We find no evidence to support the premise that the control and treatment groups had similar user conversion rates (reject the null hypothesis).

We did not notice a significant difference between group A (3.37\$) and group B (3.39\$) in terms of average expenditures. The significance level of the p value for the test of the hypothesis was 0.94%. Concluding that no significant difference in average spending between the control and treatment groups, hence the null hypothesis is not rejected.



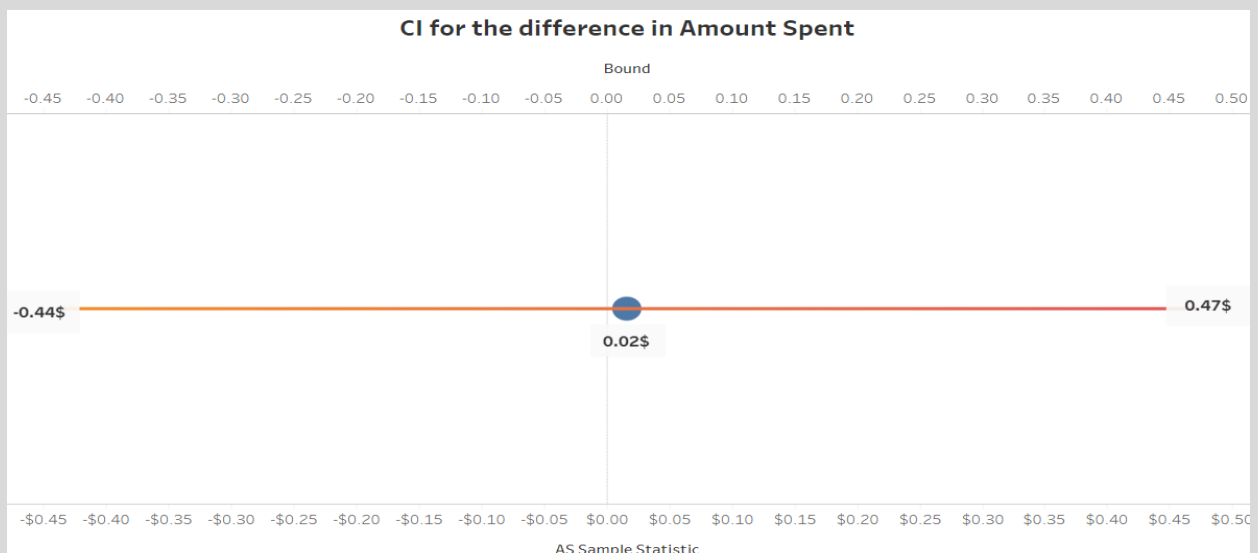
# CONFIDENCE INTERVAL FOR DIFFERENCE IN CONVERSION RATE

- The difference between the control and treatment groups, calculated at 0.71%, lies within the 95% confidence interval for conversion rates of between 0.35% and 1.07%.



# CONFIDENCE INTERVAL FOR DIFFERENCE IN AMOUNT SPENT

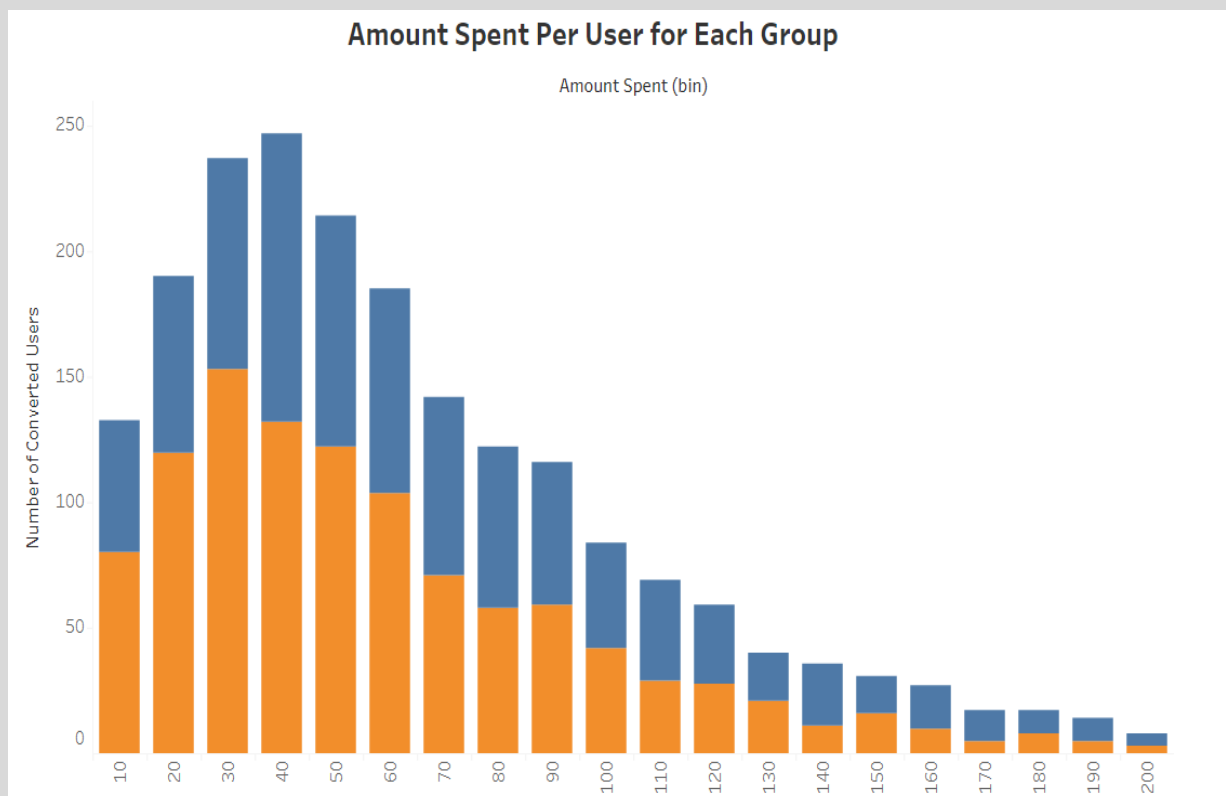
- The determined 0.02\$ difference between the control and treatment groups is within the 95% confidence interval for average amount spent of -0.44\$ to 0.47\$.



# AVERAGE AMOUNT SPENT PER USER FOR EACH GROUP

The following is apparent from the figures below:

- Majority of consumers who converted in group A had made purchases between \$40 and \$50
- Greatest proportion of group B customers who converted made purchases between \$30 and \$40
- In conclusion, consumers in the control group spent more on average than those in the treatment group, which may be due to the category and type of product they purchased.



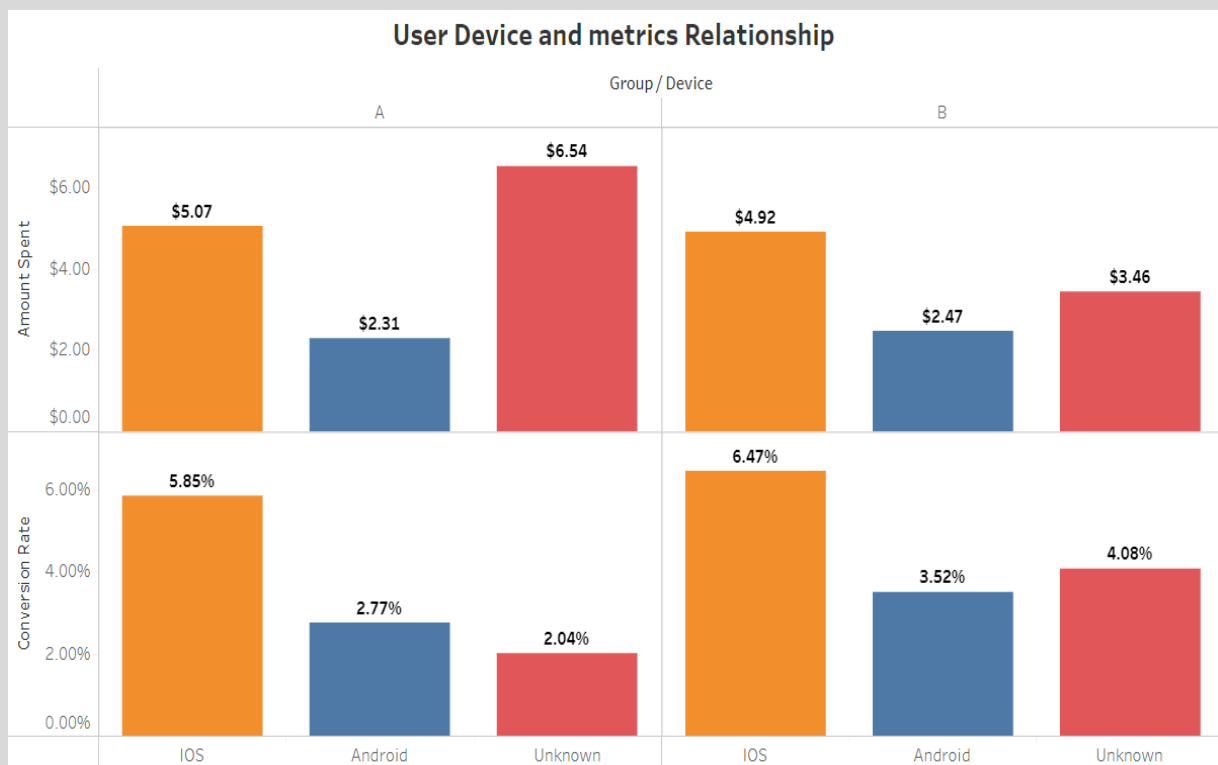
- ✓ The distributions for both groups are right-skewed (mean > mode > median), indicating that users tend to spend less on their purchases, with those who spent more than \$200 representing outliers.

# Results Analysis by Device, Gender and Country

## I. Device

The device analysis in the figure leads us to the following conclusion:

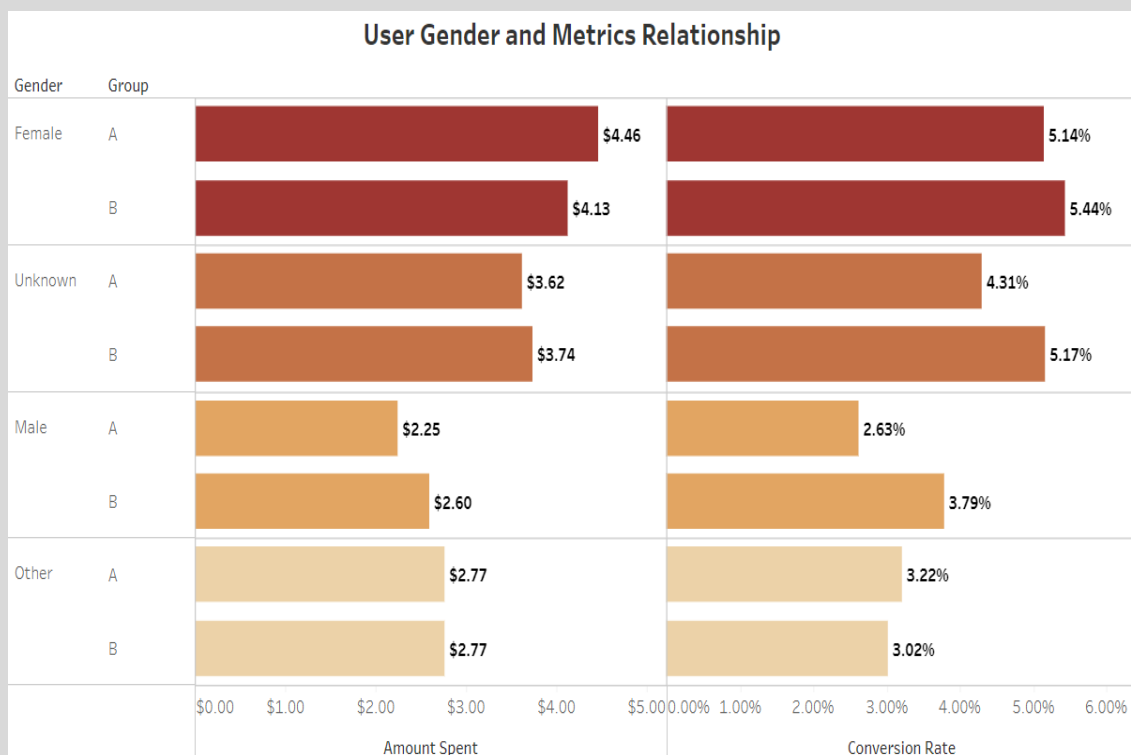
- IOS conversion rate is higher than Android users. However, the lift in conversion for Android consumers is significantly 27%.
- Consumers in the control group with IOS devices spent more than those in the treatment group, most likely due to user product interest and purchasing behavior as discussed on the previous slide.
- Users with unknown devices had a 100% lift rate despite a negative value of 47% in amount spent lifting, indicating that the treatment group performs less effectively or has a lower response rate than the control group. In other words, the intervention or treatment has a negative effect on the outcome.



## II. Gender

By analyzing the gender relationship to the metrics, the following can be determined:

- Female users have the highest conversion rate when compared to male users.
- A significant decrease in expenditures by female treatment group consumers
- Male conversion rate and amount spent lift is the highest with 44%, 15.5% respectively.
- Female consumers are more likely to convert than males.

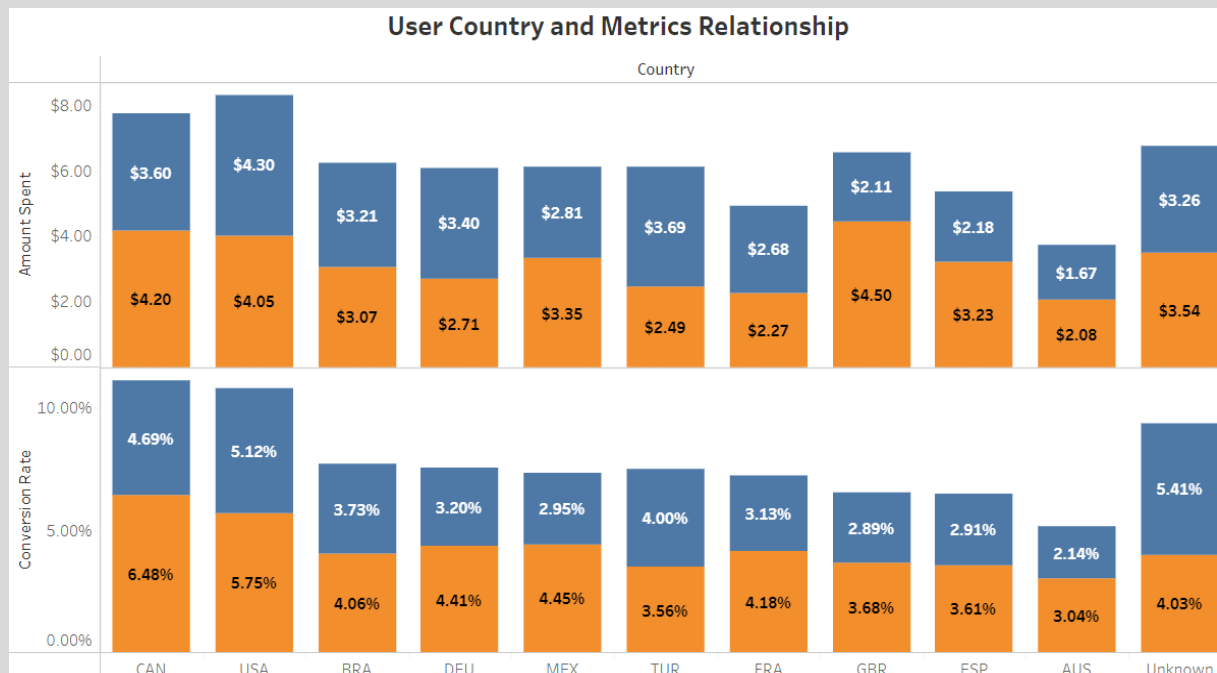


## III. Country

Monitoring country-specific metrics yielded the following results:

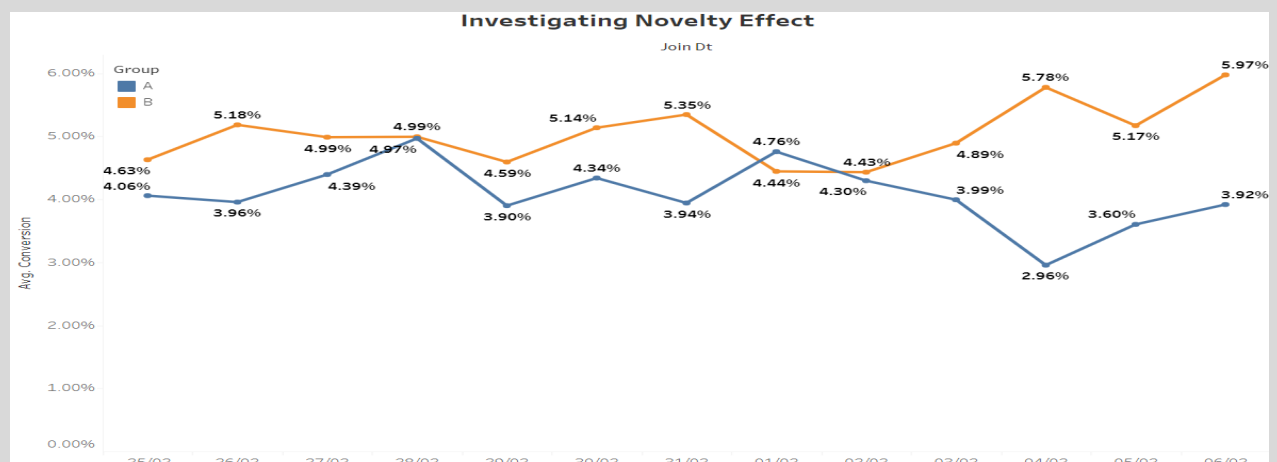
- The majority of nations saw a considerable increase in their conversion rates, with Mexico leading the way at 50.8%.
- Users from the United Kingdom performed well in terms of spending more after being exposed to the new version of the app interface, as evidenced by a 113% increase in expenditures.
- Half of nations have a negative increase in the quantity spent metric.

- The new app version was ineffective in Turkey and user's performance was subpar, registering negative values for both metrics.



## Investigating Novelty Effect

- Investigating the existence of a Novelty effect is facilitated by plotting the conversion rate against experiment duration. We observe no significant change in the direction of the linear plot and cannot discern a distinct trend, so we conclude that the user behavior was not influenced by a novelty effect





# Power Analysis

- Following the communication within the team, we have assumed that the banner launch will result in 12% of practically significant change. This will assist us in determining the minimal detectable impact given a value of 10% with regard to the requirement that MDE must be less than to practical significance.
- Using the statsig calculator with a baseline conversion rate that belongs to the control group (3.92%), a sample size of 70.9k users was determined to be conclusive. However, a sample size of 873k was determined for the amount spent by having an absolute difference equal to the control value \* MDE, which produces a value of 0.337, and a pooled standard deviation of 25.66 was computed using the pooled formula for two groups.

Metrics	Calcutaor	Values	Sample Size
Conversion Rate	Statsig	Baseline CR% = 3.92% MDE = 10%	70.9k
Amount Spent	Statulator	Absolute difference = $3.37 \times 10\%$ STDE = 25.66	873k

- Based on our research, we were able to determine that a sample size of 49K provided us with a practical change in conversion rate that was greater than 12%, registering at 18%. However, when looking at the average amount spent, the lift was only 0.5 percent, leading us to believe that the sample size we chose does not include a sufficient number of users. We believe that a larger sample would help more to visualize a meaningful shift in the amount of money spent.

# Recommendation

- ✓ After observing an improvement in a success metric represented by conversion rate, it is advised to deploy the new app version that includes the food and beverages banner.
- ✓ A statistically significant increase in conversion rate was observed.
- ✓ We cannot determine whether or not to deploy the banner if we consider sales revenues that are primarily correlated with the user's product interest and bias behavior.

## SQL Appendix

### SQL query to extract GloBox Dataset:

```
select distinct g.uid,  
               coalesce (u.country, 'No') as country,  
               coalesce(u.gender, 'Unknown') as gender,  
               g."group",  
               g.device,  
case when a.spent > 0 then 1 else 0 end as conversion,  
coalesce(cast(sum(a.spent) as numeric (10,2)),0) as amount_spent  
from users u  
left join groups g  
on u.id = g.uid  
left join activity a  
on g.uid = a.uid  
group by g.uid,2,3,4,5,6  
order by g.uid;
```

### SQL query to extract GloBox Dataset to investigate Novelty effect:

```
select distinct g.uid,  
               g.join_dt,  
               coalesce (u.country, 'No') as country,  
               coalesce(u.gender, 'Unknown') as gender,  
               g."group",  
               g.device,  
case when a.spent > 0 then 1 else 0 end as conversion,  
coalesce(cast(sum(a.spent) as numeric (10,2)),0) as amount_spent  
from users u  
left join groups g  
on u.id = g.uid  
left join activity a  
on g.uid = a.uid  
group by g.uid,2,3,4,5,6,a.spent,g.join_dt  
order by g.uid
```

## SQL query to calculate the user Conversion Rate:

```
1- SELECT ROUND (COUNT (DISTINCT uid)/ (SELECT ROUND(COUNT(id),2) FROM users) * 100,2)
AS conversion_rate
FROM activity
OR
2-select round((sum(n_actif) / round(count(u.id),2)) *100,2) AS conversion_rate
FROM
(select uid,
      count(distinct uid),
      case when spent > 0 then 1 else 0 end as n_actif
from activity
group by 1,3) sub
right join users u
on sub.uid = u.id
```

## SQL query to calculate the user Conversion Rate for each group:

### Group A:

```
select round((sum(n_actif) / (select round(count(g.uid),2) from groups g where "group" = 'A'))*100,2) as
group_a_cr
FROM
(select a.uid,
      count(distinct a.uid),
      case when a.spent > 0 then 1 else 0 end as n_actif
from activity a
left join groups g
on a.uid = g.uid
group by 1,3,g."group"
having g."group" = 'A') sub
```

### Group B:

```
select round((sum(n_actif) / (select round(count(g.uid),2) from groups g where "group" = 'B'))*100,2) as
group_a_cr
FROM
(select a.uid,
      count (distinct a.uid),
      case when a.spent > 0 then 1 else 0 end as n_actif
from activity a
left join groups g
on a.uid = g.uid
group by 1,3,g."group"
having g."group" = 'B') sub
```

## SQL query to calculate the user Average amount spent for each group

```
with r as
(select sum(spent) as sum_ga
from activity a
right join groups g
on a.uid = g.uid
group by g."group"
having g."group" = 'A'),
b as
(select sum(spent) as sum_gb
```

```
from activity a
right join groups g
on a.uid = g.uid
group by g."group"
having g."group" = 'B')
```

```
select r.sum_ga / (select round(count(g.uid),2) from groups g where "group" = 'A')
```

```
from r
union
select b.sum_gb / (select round(count(g.uid),2) from groups g where "group" = 'B')
from b
```

## URLs:

Spreadsheet containing statistical analysis for hypothesis testing and 95% confidence interval

[https://docs.google.com/spreadsheets/d/1QXULeQLZJfD\\_nTECG7\\_VREsxb9X1Lk92St76OPhIrow/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1QXULeQLZJfD_nTECG7_VREsxb9X1Lk92St76OPhIrow/edit?usp=sharing)

Tableau illustrating the different plots for the analysis:

<https://public.tableau.com/app/profile/djamel.fitas/viz/GloBoxABTestingAnalysis/GloBoxAnalysis>