GloBox A/B Test Results

James Franchino - Data Analyst

Introduction

I was given the task of analyzing an A/B test run by an e-commerce company called GloBox. The test was run on the mobile website, and it involved showing a banner at the top of the page to the test group, while the control group did not see the banner. The goal was to increase revenue from the food and drink category. My task was to analyze the results of the A/B test and provide a recommendation on whether GloBox should launch the experience to all users.

Group A: Control

existing landing page

Group B: Treatment

landing page with food & drink banner

Methods

The data was stored on a remote database (link to data). I used SQL to combine the relevant data into one table which I then downloaded as a .CSV

```sql
SELECT
  u.id AS user_id,
  u.country,
  u.gender,
  g.group,
  COUNT(DISTINCT a.dt) AS num_purchase_days,
  SUM(a.spent) AS total_spent,
  CASE
    WHEN SUM(CASE WHEN a.spent > 0 THEN 1 ELSE 0 END) > 0 THEN 1
    ELSE 0
  END AS converted
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.group;
```

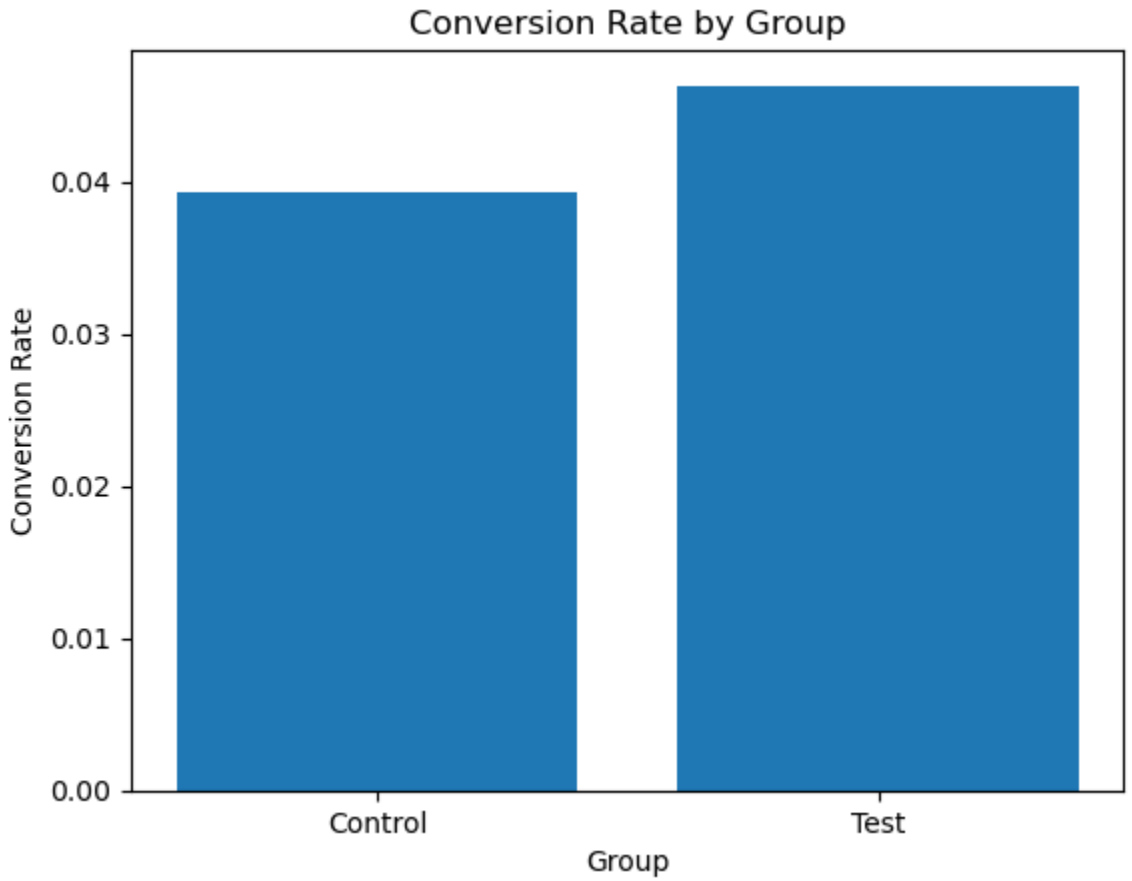
I used Python and various statistical analysis methods to perform my analysis. I first cleaned the data by filling in null values and converting the 'converted' column to a boolean.

Test Results

I then tested the null hypothesis that there is no difference in conversion rates between the control and test groups. I used a two-sample t-test to compare the mean conversion rates between the groups. I found a p-value of 0.0001113, which is less than the significance level of 0.05, and therefore rejected the null hypothesis. This means there is a significant difference in

conversion rates between the two groups. The conversion rate for the control group was 3.92%, while for the test group it was 4.63%.

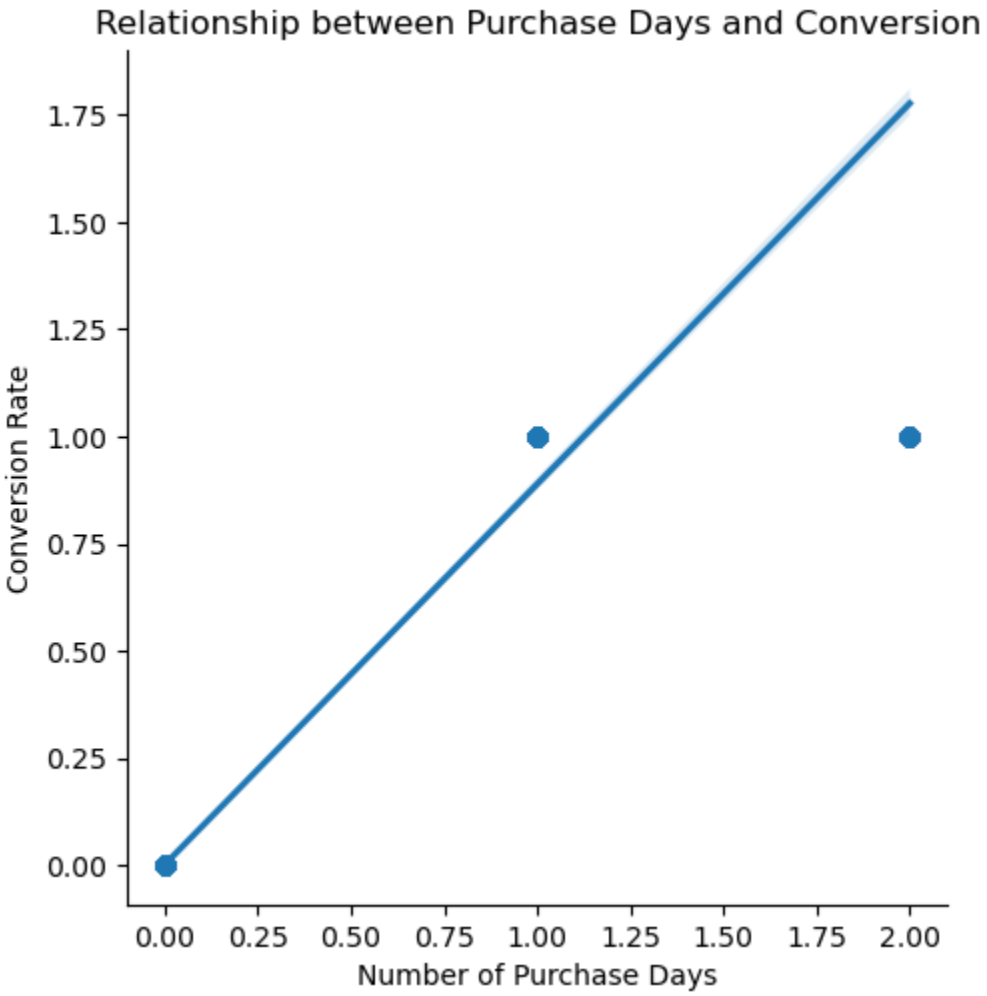
I also calculated the 95% confidence interval for the conversion rates of both groups. For the control group, I used the normal distribution and found a confidence interval of (0.0387, 0.0492). For the test group, I also used the normal distribution and found a confidence interval of (0.0451, 0.0576). Finally, I calculated the 95% confidence interval for the difference in conversion rates between the test and control groups. using the normal distribution and unpooled proportions for the standard error. The confidence interval was (0.003486, 0.010654).



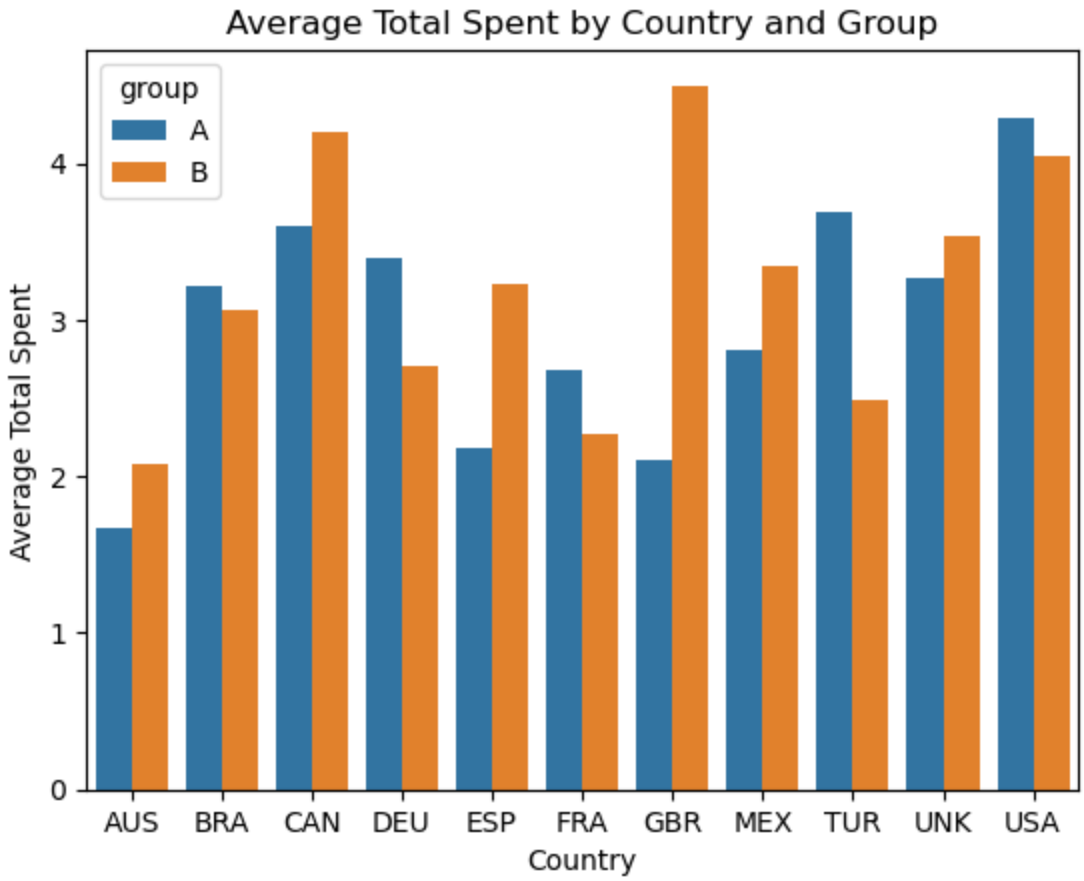
I next tested the null hypothesis that there is no statistically significant difference in the total amount spent between the two groups. Using a two-sample t-test and found a p-value of 0.9439, which is greater than the significance level of 0.05, and therefore failed to reject the null hypothesis. This means there is no significant difference in the total amount spent between the two groups. The average amount spent per user in the control group was $3.375, while in the test group it was $3.391. I also calculated the 95% confidence interval for the average amount spent per user in both groups. For the control group, I used the t-distribution and found a

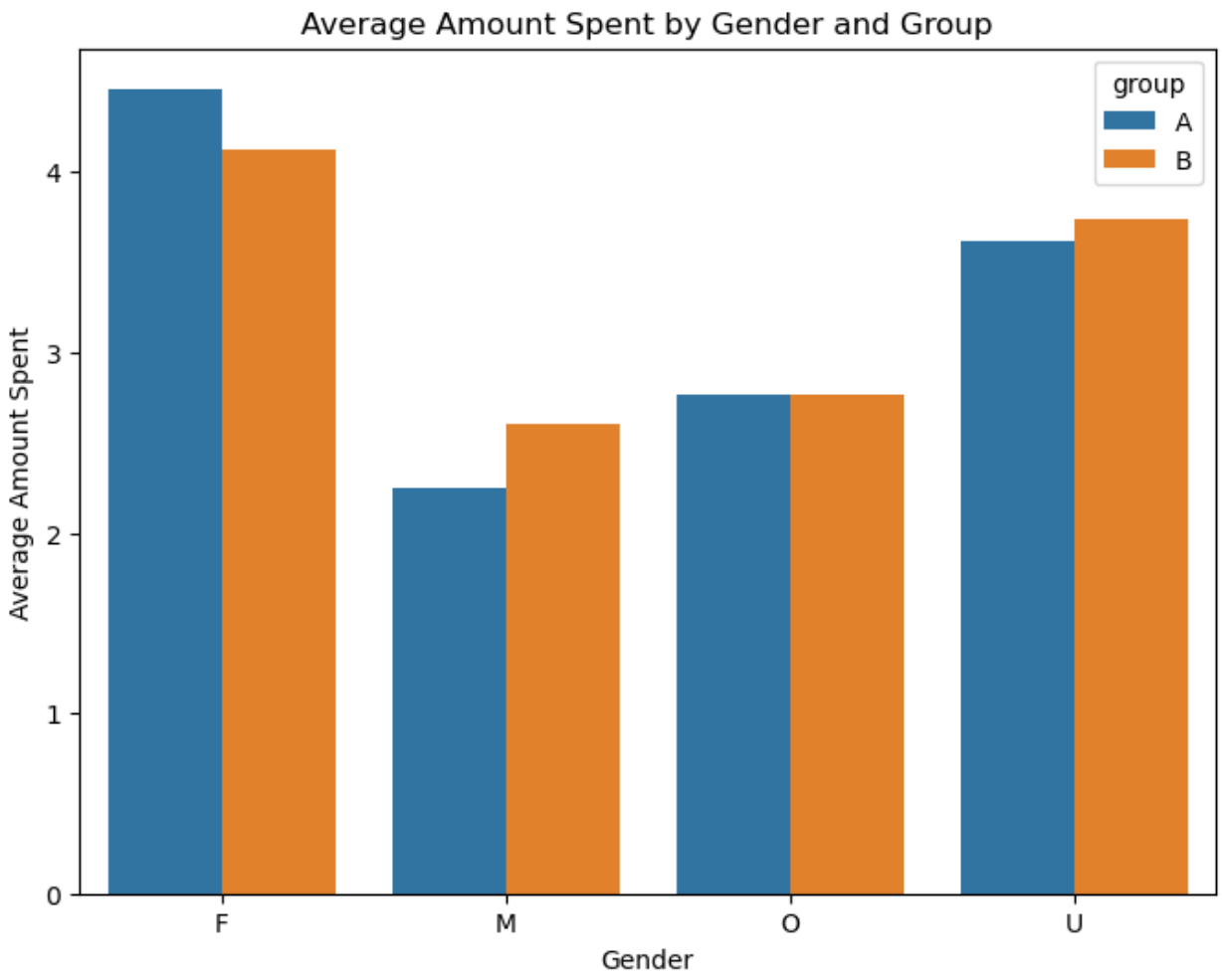
confidence interval of ($3.327, $3.424). For the test group, I also used the t-distribution and found a confidence interval of ($3.335, $3.448).

I looked at the correlation between the number of purchase days and conversion. Finding a Pearson correlation coefficient of 0.9727, which indicates a strong positive correlation. I rejected the null hypothesis that there is no correlation between the number of purchase days and conversion. I calculated the 95% confidence interval for the correlation coefficient, which was (0.9722, 0.9732).



I finished by visualizing the amount spent per group by country, amount spent per group by gender, and gender breakdown by country. Finding that the UK had the highest average amount spent per user in both groups and that females in the control group spent more than males in both groups. I found that the highest average amount spent per user by gender and country was by UK females.


Average Total Spent by Country and Group
group
A
B
Average Total Spent
0
1
2
3
4
AUS
BRA
CAN
DEU
ESP
FRA
GBR
MEX
TUR
UNK
USA
Country



Conclusion

In conclusion, this analysis shows that there is a significant difference in conversion rates between the control and test groups, and that there is a strong positive correlation between the number of purchase days and conversion.

Based on the analysis I have conducted, there is strong evidence to suggest that the treatment group (test group) had a higher conversion rate than the control group. This is supported by the results of our hypothesis test and confidence interval calculation. Additionally, the 95% confidence interval for the difference in conversion rates between the test and control group does not contain 0, further indicating that there is a statistically significant difference between the two groups.
However, I did not find a statistically significant difference in the total amount spent between the two groups. The 95% confidence interval for the difference in average amount spent per user also contains 0. Therefore, it is unclear if the treatment group resulted in a significant increase in revenue.

I found that there is a strong positive correlation between the number of days a user is in the experiment and their likelihood of conversion. This is not surprising as users who spend more time on the site have more opportunities to make a purchase.
Lastly, I looked at the breakdown of conversion rates and average amount spent per user by country and gender. While there are some differences between groups, none of them appear to be significant enough to make a recommendation based on them alone.
Taking all of these results into consideration, I recommend launching the newer version of the site to all users. The increase in conversion rate observed in the test group suggests that the changes made to the site were effective in driving more purchases, even if the total amount spent per user did not increase significantly. However, I also recommend monitoring the revenue over time to ensure that there are no negative impacts from the changes.