GloBox Food & Drink A/B Test

Increasing Revenue and Raising Awareness for an Expanding Product Category

Company

- GloBox is an e-commerce company specializing in boutique fashion items, high-end decor products, and increasingly, food and drink offerings.
- GloBox is well-known for sourcing unique products from around the world and providing an exceptional shopping experience.

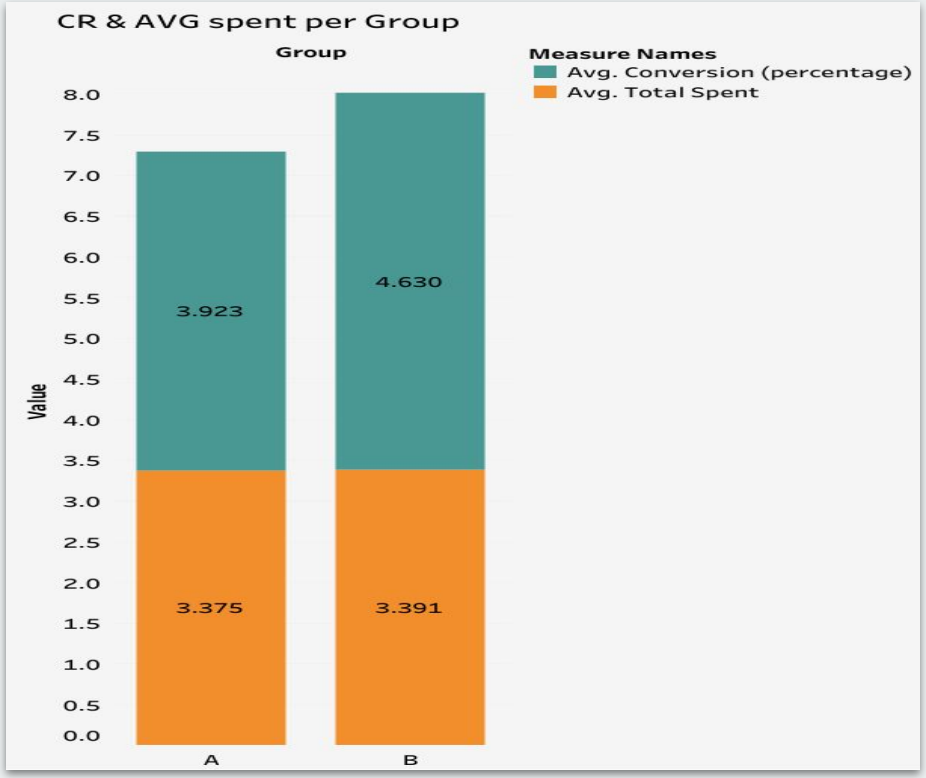
Context

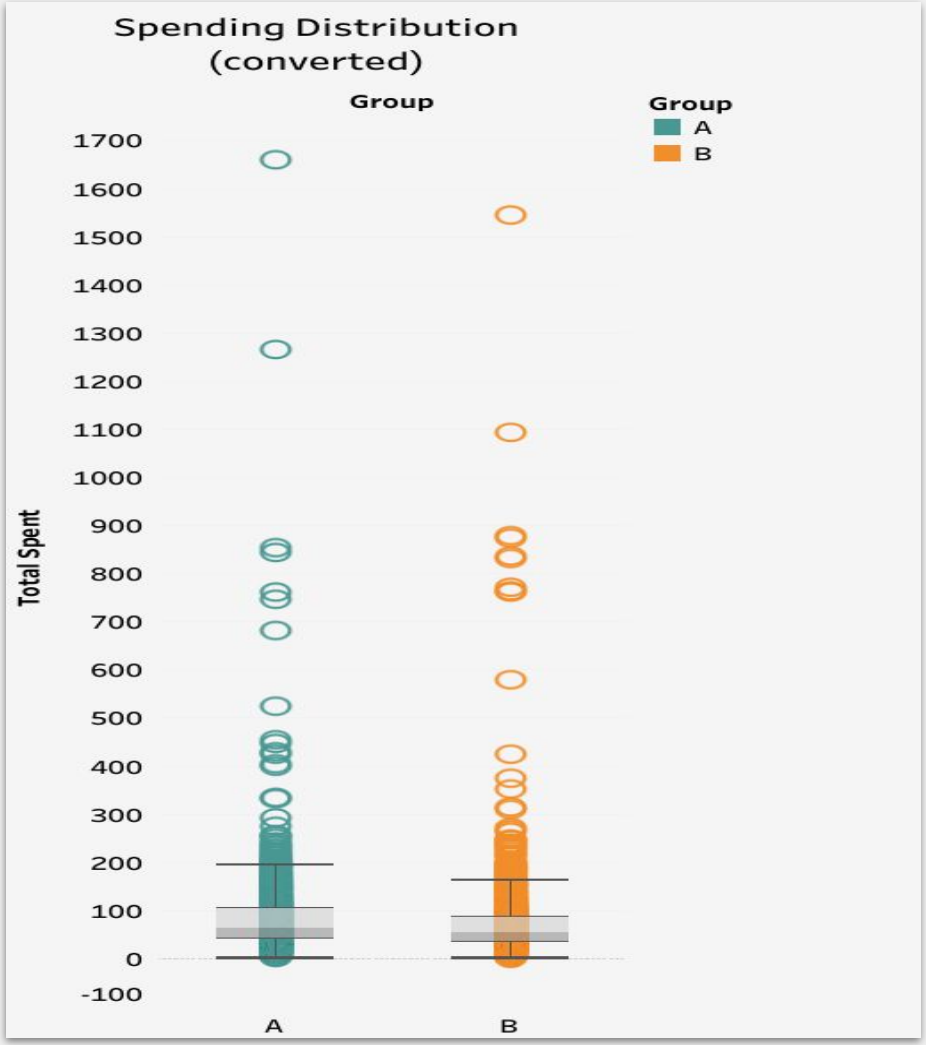
- GloBox's recent growth in food and drink products requires increased revenue and awareness. The A/B test is a means of determining the most effective promotion and sales strategies to showcase their offerings to customers.

Problem statement

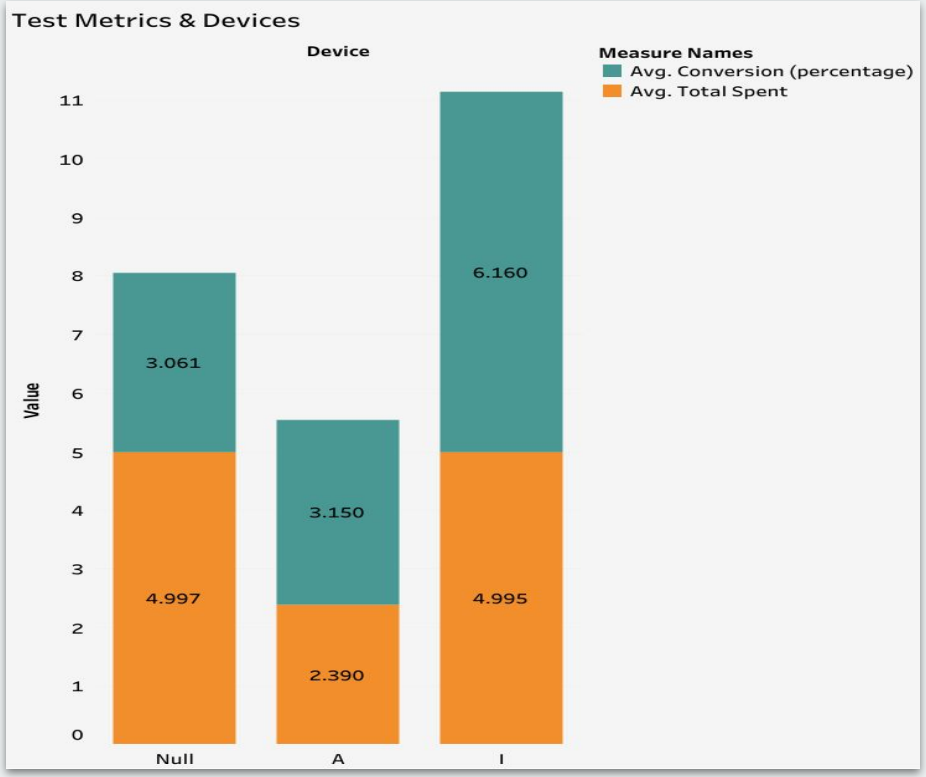
- The primary objective of the A/B test was to determine whether the implementation of a banner showcasing key food and drink products on the GloBox mobile website would positively influence user behavior and lead to increased conversions and total amount spent per user.

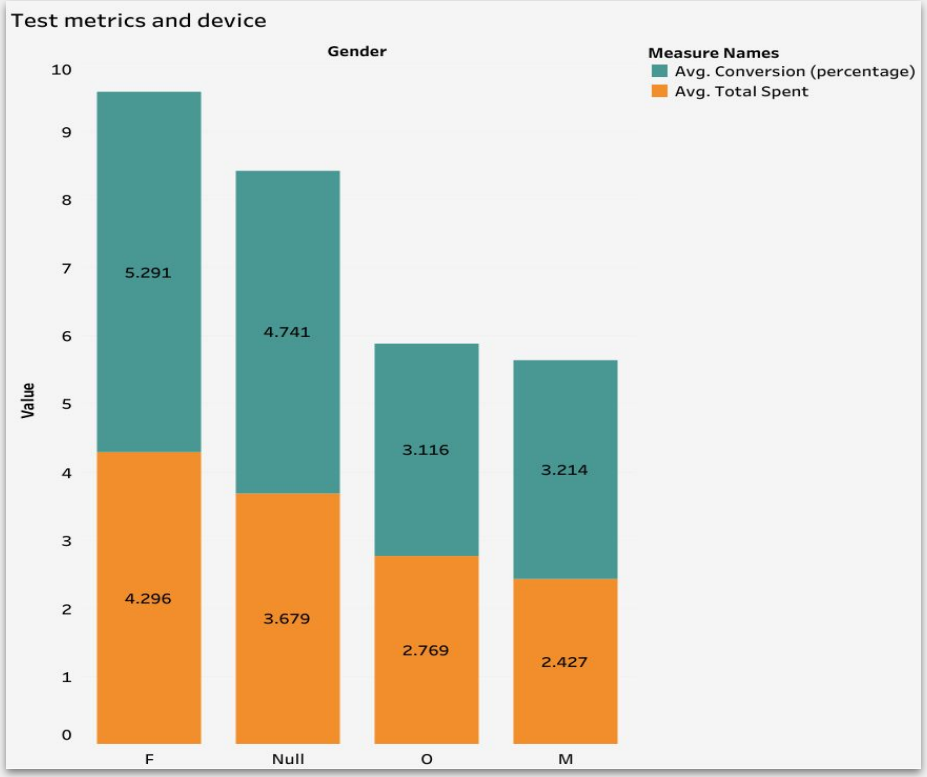
Exploratory Data Analysis



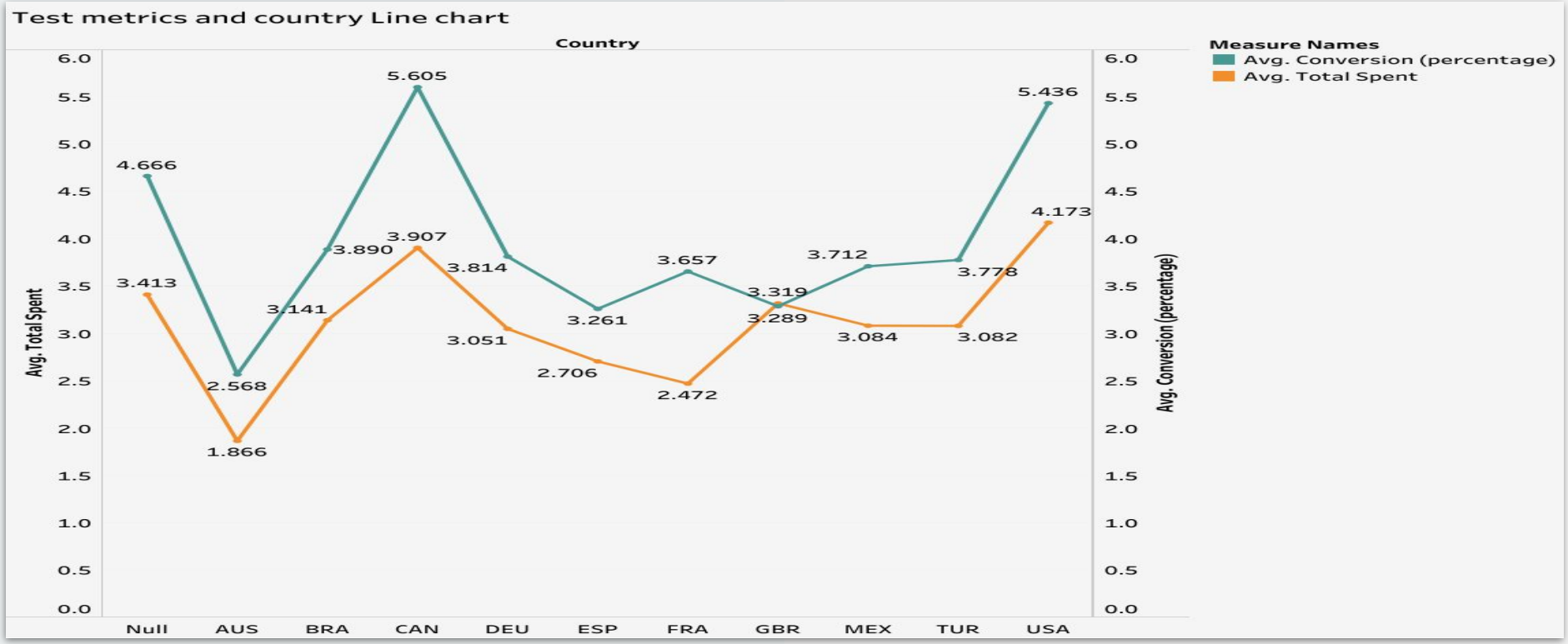


Exploratory Data Analysis

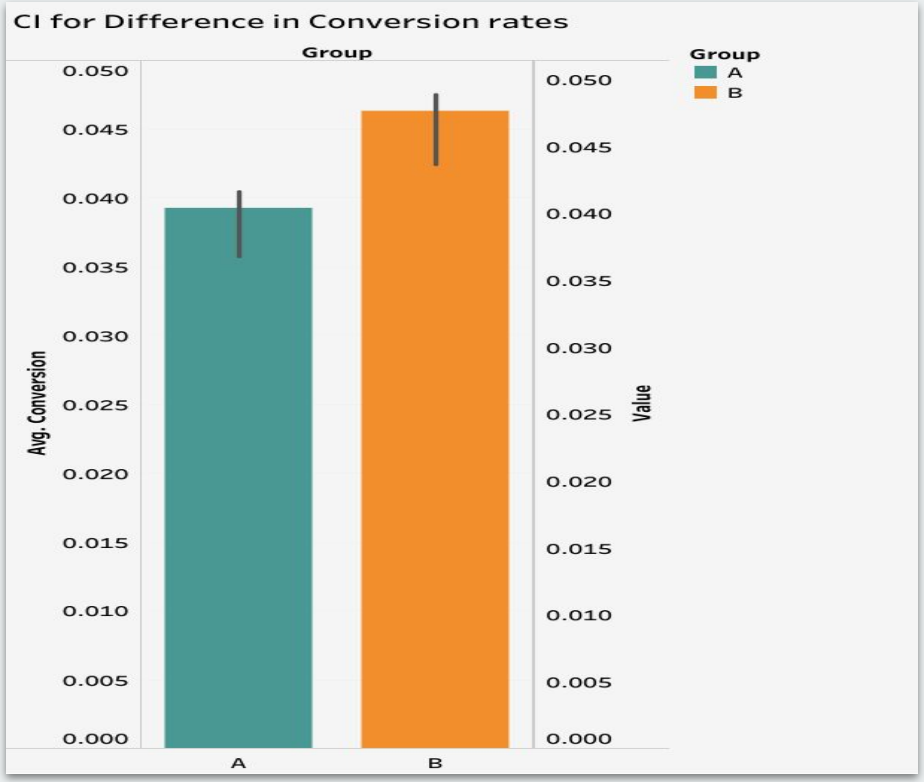




Exploratory Data Analysis

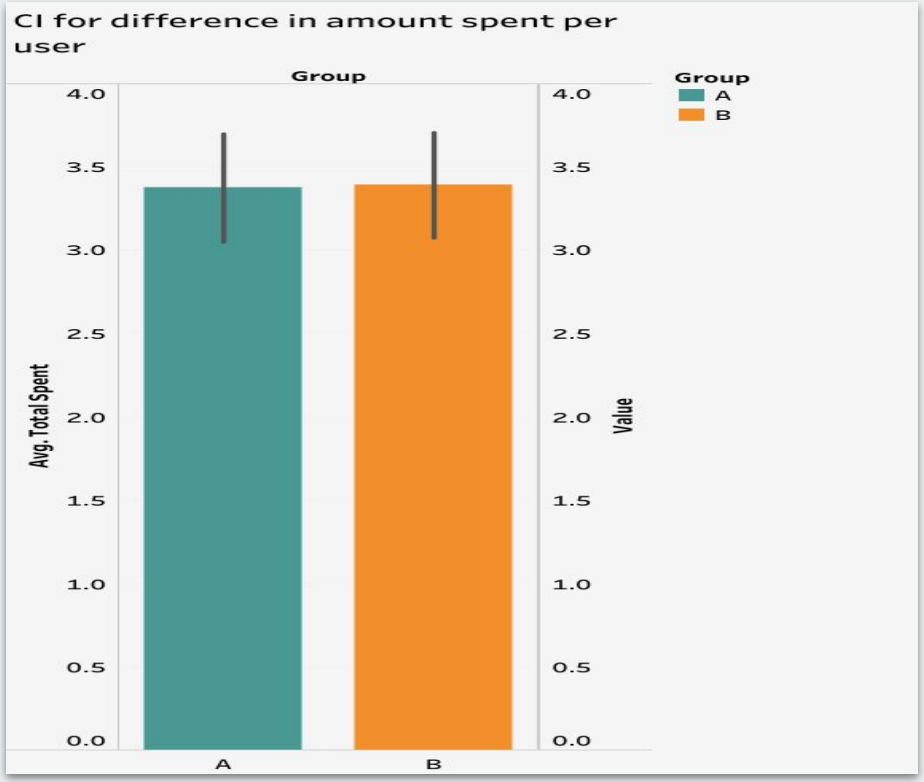


95%Confidence intervals for difference in conversion rates between group A & B, using **pooled proportion for standard error.**



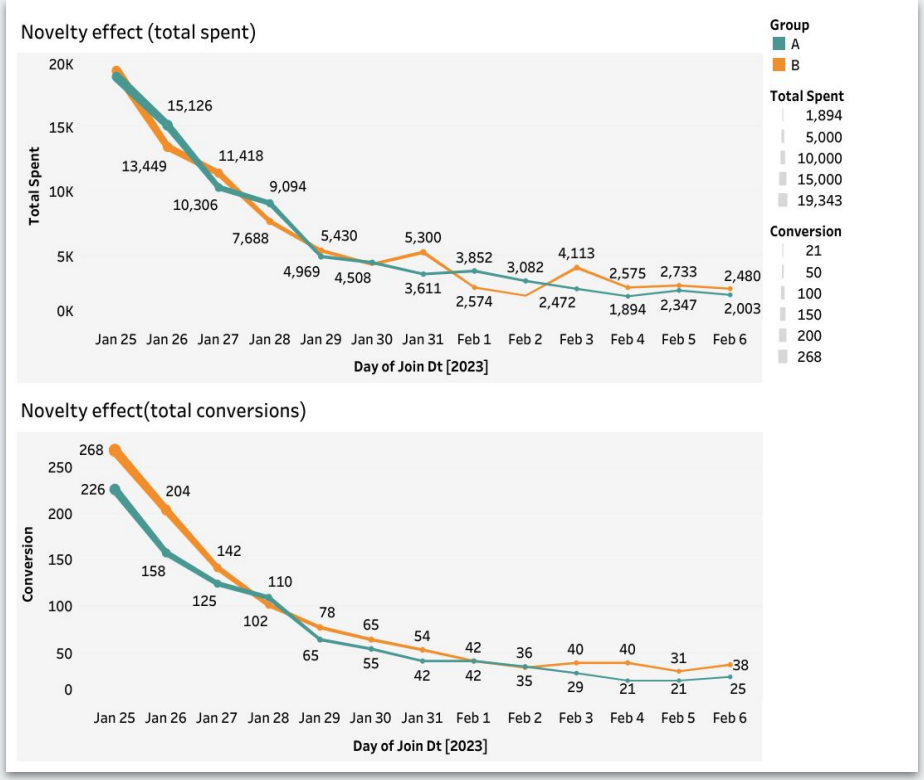
- The 95% confidence intervals for Group A and Group B do not overlap, indicating a statistically significant difference between the two groups at the 95% confidence level.
- This lack of overlap signifies there is a less than 5% probability that the observed difference in conversion rates is due to chance alone, and therefore the difference is likely indicative of an actual effect present in the underlying population conversion rates, rather than a chance occurrence.

95% Confidence intervals for difference in average spent amount per user between group A & B, assuming **T-distribution and unequal variance**



- The 95% confidence intervals for the difference in average spent amount per user between group A and group B overlap, it suggests that there is no statistically significant difference between the two groups at the 95% confidence level.
- The larger the overlap, the less likely it is that the difference represents a real effect rather than random variation in the samples. With very large overlap, it is highly probable that the difference is due to chance.

Novelty Effect



- In A/B testing, if one test group experiences a novelty effect, their metrics may initially spike simply due to the novelty of the element being tested, rather than it truly driving better performance.
- The lack of a significant difference in the trends of the lines for total conversion and total amount spent per group over time suggests that there is no evidence of a significant novelty effect.

Limitation

- **Sample size imbalance:** The sample sizes for Group A and Group B are quite large, but there is still a slight imbalance. This could potentially bias the results, especially for the confidence intervals. It would be better to have exactly equal sample sizes for the two groups if possible.
- **Multiple testing issues:** Conducting multiple tests (4 in this case) on the same data increases the chances of false positive results (finding a significant difference when there is none). Some adjustment to the alpha levels, like a Bonferroni correction, may be needed.
- **Possible biases:** There are many possible biases that could influence the results, including selection bias, response bias, measurement bias, observer bias, etc. The study design and procedures should be evaluated to minimize the impact of biases.
- **Assumption violations**: The tests assume certain conditions, like normality and independent samples. If these assumptions are violated, the results may not be valid. The data and residuals should be checked to ensure the assumptions are met.

Recommendation

- The hypothesis test for the difference in conversion rates between Group A and Group B showed a p-value of 0.0001114119853. This is very statistically significant and shows strong evidence that the conversion rates between the two groups are different. Since the confidence intervals do not overlap, this further confirms the difference.
- However, the hypothesis test for the difference in average spend per user showed a p-value of 0.9438557532. This indicates no statistically significant difference between the two groups. The confidence intervals also overlap, further confirming no meaningful difference.
- **It is recommended not to proceed with the launch of the new banner feature based on total amount spent per group. While there were significant differences in conversion rates between the two groups, it was observed that there was no significant difference in the average amount spent per group. Therefore, it is advised to explore alternative strategies that may better align with the overall goals and objectives.**