# A/B Test Analysis

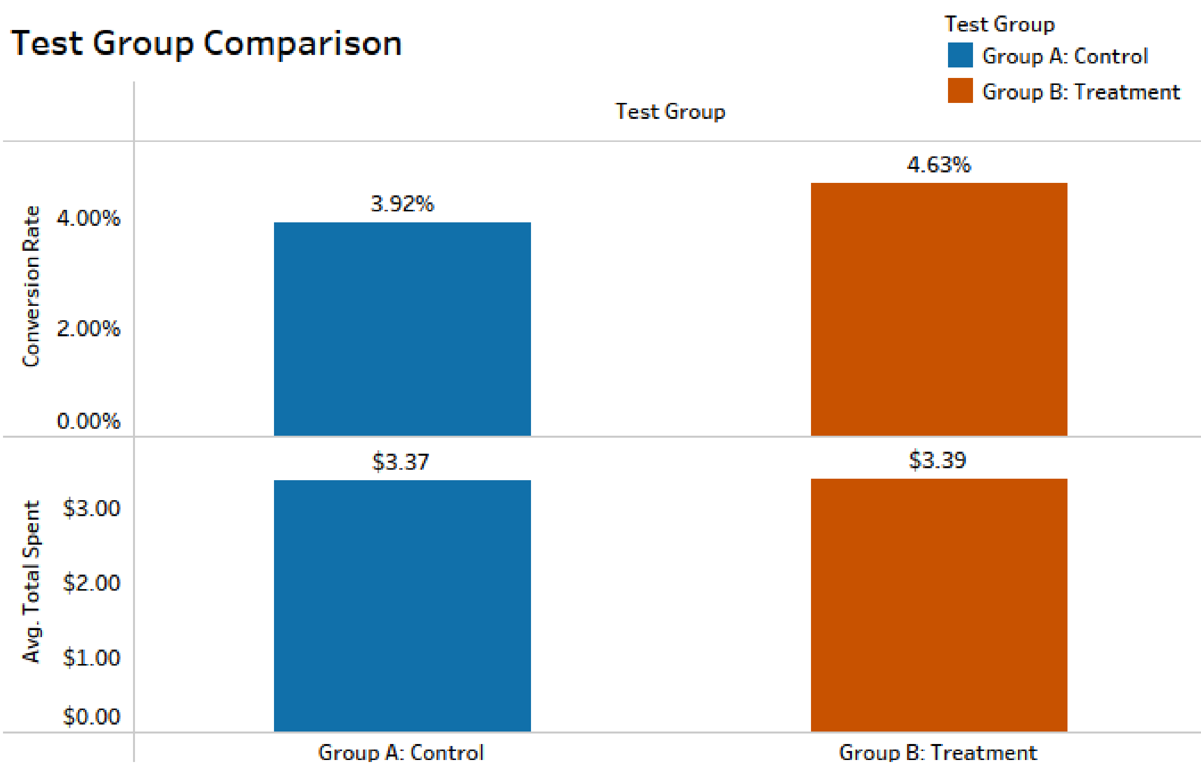
Globox Food and Drink Banner Test

Charu Kesarwani

# Table of Contents

# Overview

Having conducted a thorough analysis of the A/B test outcomes and considering the supplementary insights, I am confident in **recommending the launch of the banner**. The A/B test yielded a substantial discrepancy in conversion rates between the control and treatment groups, strongly indicating the positive effect of the banner on driving conversions.




By launching the banner, we can expect to significantly heighten awareness within the food and drink category, leading to an increase in conversions. It's important to note, however, that no conclusive evidence has emerged regarding differences in the average total spend.

In light of this, I suggest maintaining a vigilant approach to tracking average spending trends. To gain more conclusive insights into potential mean variations, it might be worthwhile to consider conducting further experiments or expanding the sample size. Doing so will provide a deeper understanding of how the banner impacts user spending behavior, offering valuable insights to guide future decision-making.

# Exploring Methodology and Data Analysis

In this segment, we explore the structured methodology and methodologies employed to thoroughly analyze the outcomes of the A/B test.

## Step 1 - Gathering Data with SQL:

To commence, I extracted essential data for the A/B test from our ecommerce platform's database. I collaborated with three crucial tables.

Users Table: This table contains information about users, like their IDs and demographic details such as country and gender.

Groups Table: Here, I found records of which users were assigned to which A/B test group. It includes details such as user IDs, test group assignments, enrollment dates, and device types (I for iOS, A for Android).

Activity Table: This table tracks user purchase activity, including user IDs, purchase dates, device types, and the amount spent in USD.

I used SQL queries to combine information from these tables and gather relevant insights. This included analyzing user interactions, conversion rates, and average spending for both the control and treatment groups.

## Step 2 - Testing Hypotheses with Spreadsheets:

With the data collected, I moved on to hypothesis testing. This process aimed to compare conversion rates and average spending between Group A (control) and Group B (treatment). I set up two hypotheses:

Null Hypothesis ($H_0$): There is no significant difference in conversion rates and average spending between the two groups.

Alternative Hypothesis ($H_1$): There is a difference in conversion rates and average spending between the two groups.

I conducted both the t-test and z-test at a significance level of 5% using Google Spreadsheets. These tests helped me determine the statistical significance of the observed differences.

## Step 3 - Visualizing Data with Tableau:

Utilizing Tableau for data visualization, I probed into user behavior and spending patterns. The visualizations highlighted:

Distribution of Average Spending: Analyzed spending distribution, identifying notable patterns.

Gender Analysis: Explored gender-based variations in engagement and conversion rates.

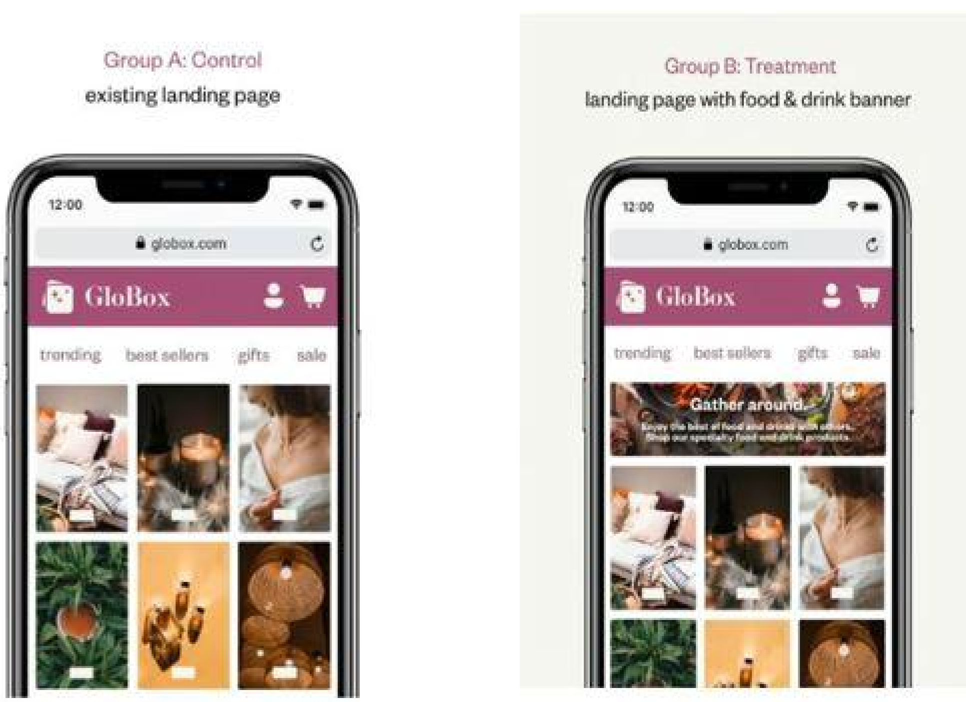
Device Analysis: Investigated user trends between iOS and Android devices.

Geographical Patterns: Examined user behavior across different countries and regions

# Context

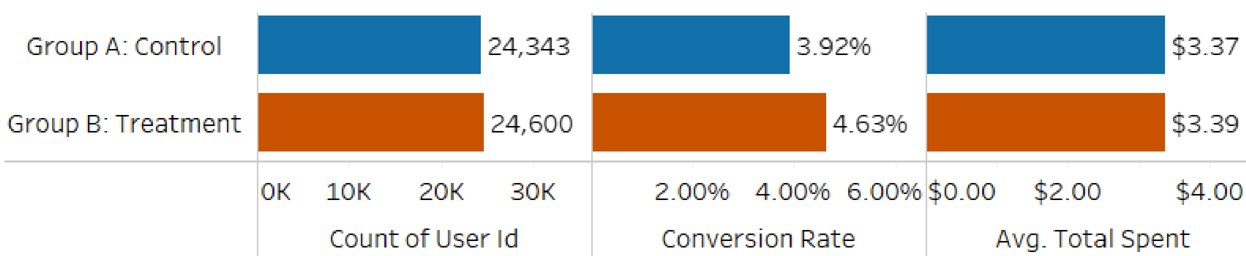
GloBox, a prominent player in the e-commerce arena, renowned for its offerings of distinct and high-quality products, embarked on a significant A/B test aimed at boosting awareness and revenue within their food and drink category. This strategic endeavor unfolded exclusively on their mobile website, with a direct focus on users navigating to the GloBox main page. Participants were divided into two groups, the Control and Treatment groups, through a random assignment process linked to their enrollment dates. Within the Treatment group, an attention-grabbing banner, showcasing the array of food and drink products, was prominently displayed. Conversely, the Control group did not encounter this visual cue. The central objective of the study was to meticulously monitor user conversions, which materialized through a user's purchase activity, whether occurring on the very day of enrollment or in subsequent instances.

Unveiling the layers of this dataset offered a window into understanding the reverberations of the food and drink banner on the intricate landscape of conversion rates, intricately tied to user expenditure dynamics. Beyond these explorations, the dataset unfolded a tapestry of insights into an array of user behaviors, informed by factors like gender, device preference, and geographic location. These insights bear the potential of guiding lights, steering the calibration of marketing strategies that resonate more effectively with distinct user segments. As we embark on this empirical journey, we acknowledge its role as a guiding compass for strategic decision-making, underscoring its significance as a foundational cornerstone in comprehensive comprehension.



# Results

## Inferential Statistics:

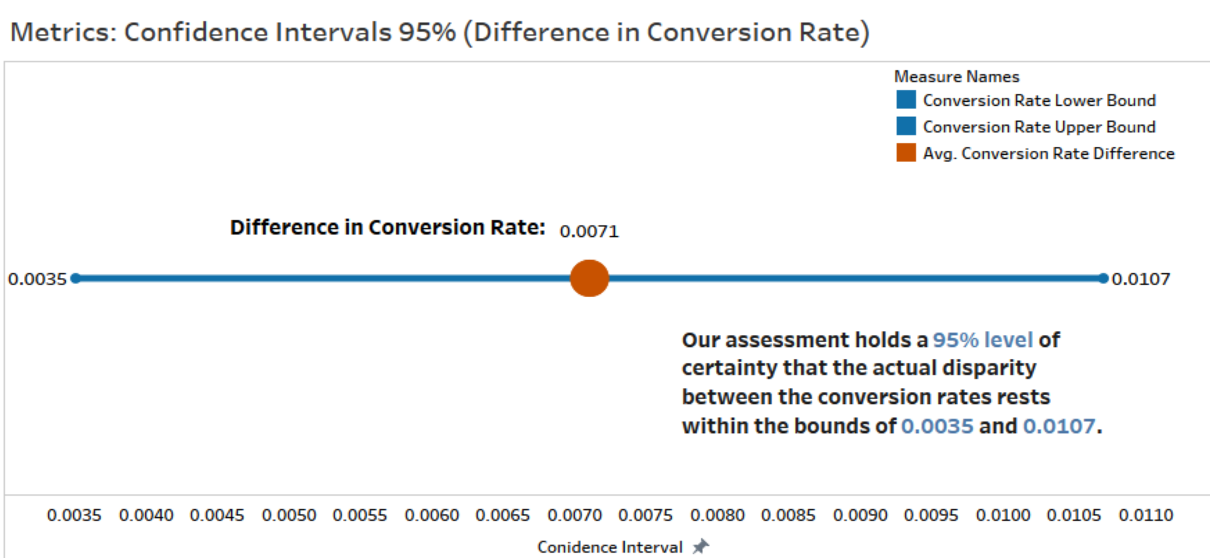



### Conversion Rate Analysis:

We wanted to see if there was a difference in conversion rates between Group A and Group B. Our null hypothesis ($H_0$) suggested that there is no difference, while our alternative hypothesis ($H_1$) indicated that there is a difference.

We chose a significance level of 5%, and the calculated p-value turned out to be 0.0001. The 95% confidence interval for the difference in conversion rates was found to be 0.0035 to 0.0107.

Based on our analysis, we rejected the null hypothesis. This means that there is indeed a significant difference in conversion rates between the two groups. The confidence interval suggests that the treatment group probably has a higher conversion rate compared to the control group. However, remember that the 95% confidence interval means that in about 95 out of 100 similar experiments, the true difference in conversion rates would fall between 0.35% and 1.07%.

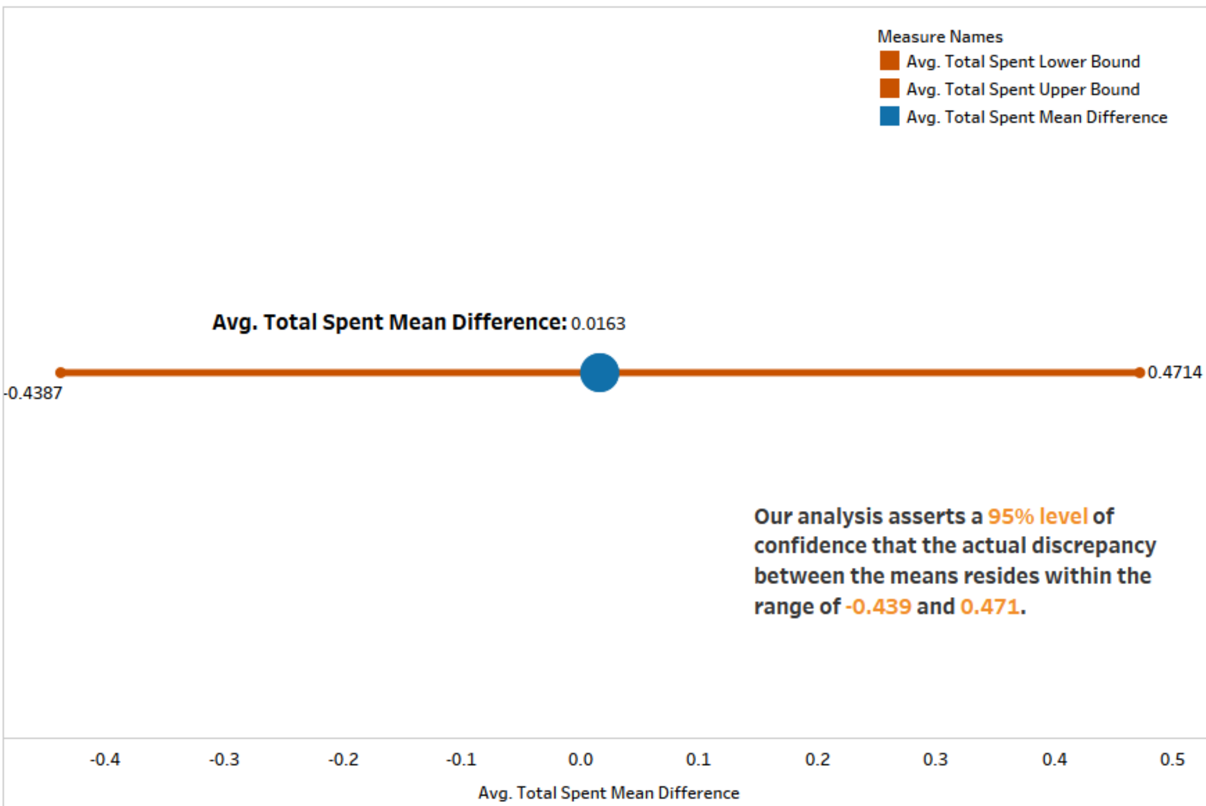


## Average Amount Spent Analysis:

We also looked into the average amount spent per user in both groups. The null hypothesis ($H_0$) stated that there's no difference, while the alternative hypothesis ($H_1$) suggested otherwise.

The p-value from our t-test was 0.9443, and the confidence interval ranged from -0.4387 to 0.4714. Based on our analysis, we couldn't reject the null hypothesis. This means that there's no statistical difference in the average amount spent per user between the two groups. The confidence interval containing zero supports this conclusion.

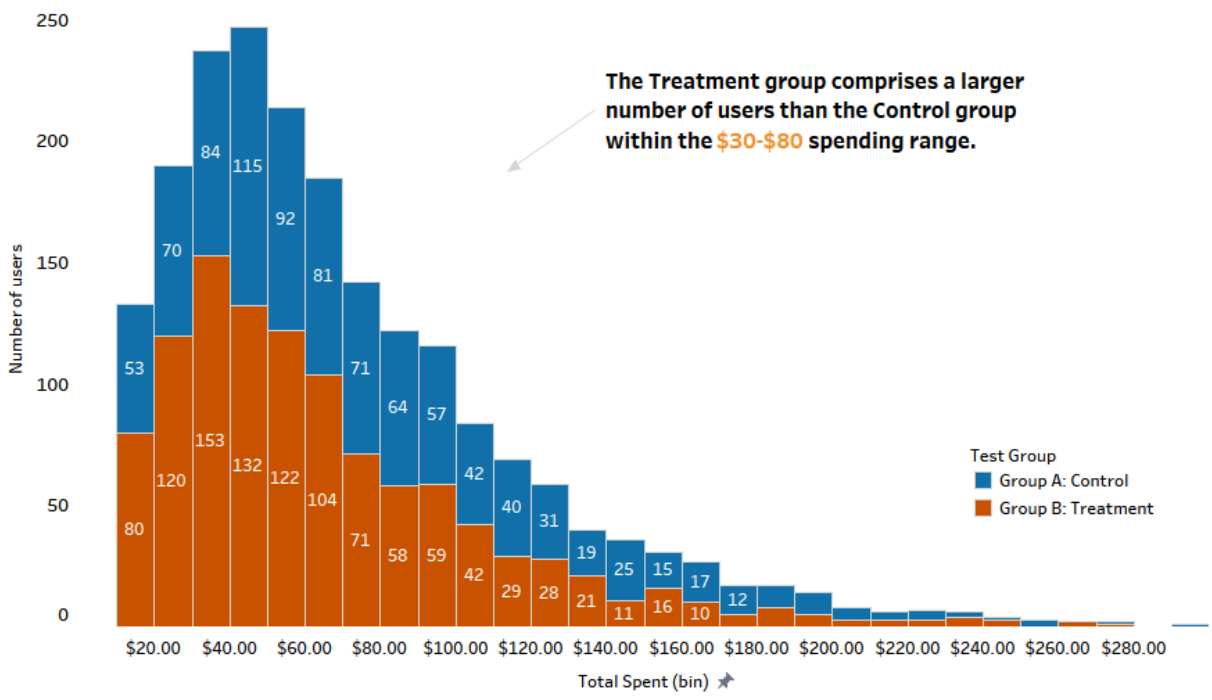
Metrics: Confidence Intervals 95% (Difference in Average Total Spent)



## Distribution of Average Total Spent:

### Observations:

In the Control Group, 115 users displayed spending between $40 to $50, while the Treatment Group had 153 users spending $30 to $40. This variation suggests the food and drink banner might influence user spending. The Control Group's higher spending might stem from other factors, whereas the Treatment Group's lower spending could reflect a conscious response to the banner. This underlines the banner's potential impact on user behavior and its significance in shaping effective marketing strategies.




### Suggestion:

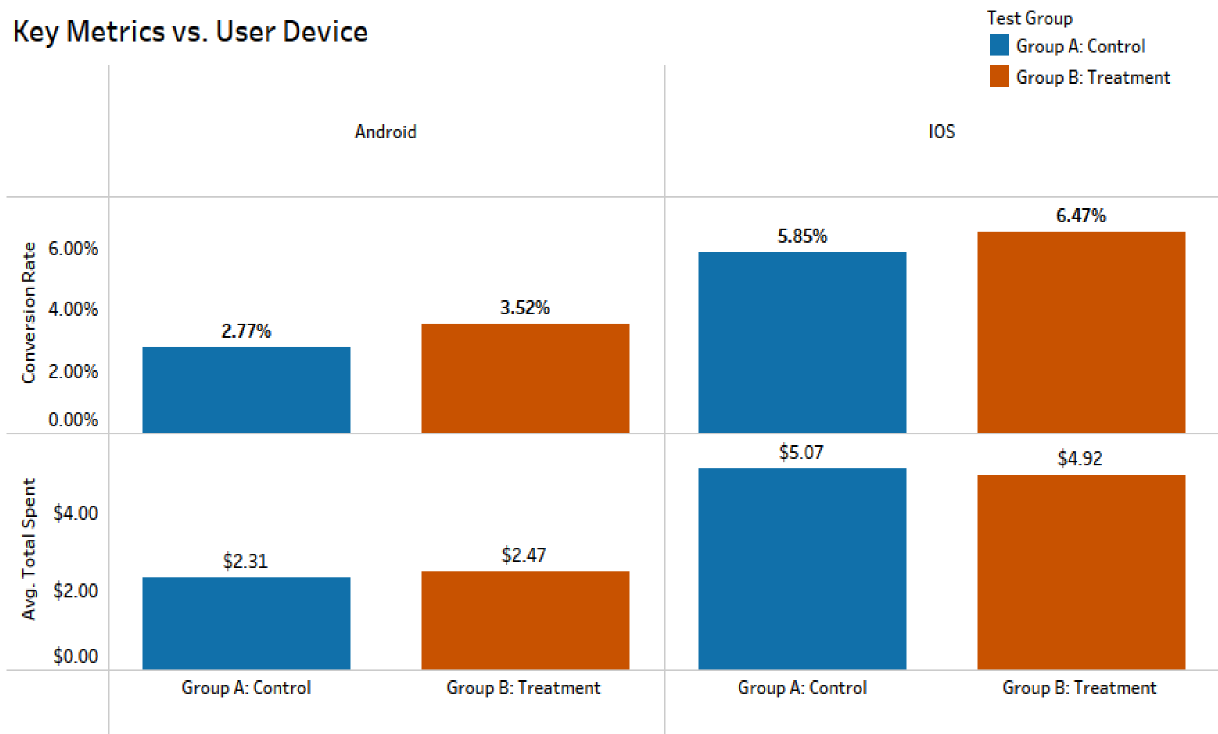
A series of strategic suggestions can be employed to enhance the effectiveness of your approach. Begin by conducting an in-depth investigation into the underlying drivers of spending behaviors in both groups. This exploration should specifically focus on understanding the factors responsible for the prevailing spending range of $40-$50 in Group A and $30-$40 in Group B. Following this, a thorough analysis of pricing strategies, promotions, and product features within these price brackets will illuminate the effectiveness of your current strategies. Building on these insights, consider tailoring your marketing messages and campaigns to resonate with users, aligning with their distinct spending preferences. To ensure a dynamic and adaptable strategy, consistently monitor spending patterns over time and make necessary adjustments to maintain alignment with evolving user behaviors.

## Analysis by Device:

### Observations:

On the Android front, the Treatment Group exhibited a remarkable surge of about 28.78% in conversion rates among Android users, indicating the banner's effectiveness. Although initial conversion rates and average spending were relatively lower for Android users, a moderate 7.91% increase in average spending coupled with a positive uptick in conversion rates signals positive potential.

The introduction of the banner yielded encouraging results, particularly for iOS users within the Treatment Group. This subgroup witnessed a noteworthy boost of approximately 11.40% in conversion rates following banner exposure. While this improvement is promising, it's important to note a slight trade-off as the average spending among iOS users decreased by around 2.22% post-banner implementation.




### Suggestions:

To harness the increased conversion rate among iOS users, the Growth team can focus on streamlining the iOS conversion journey. This involves iterative A/B testing of critical touch points such as the checkout process and product recommendations. Tailoring sales strategies to user preferences and introducing targeted promotions for higher-value items could encourage elevated spending per transaction among iOS users.

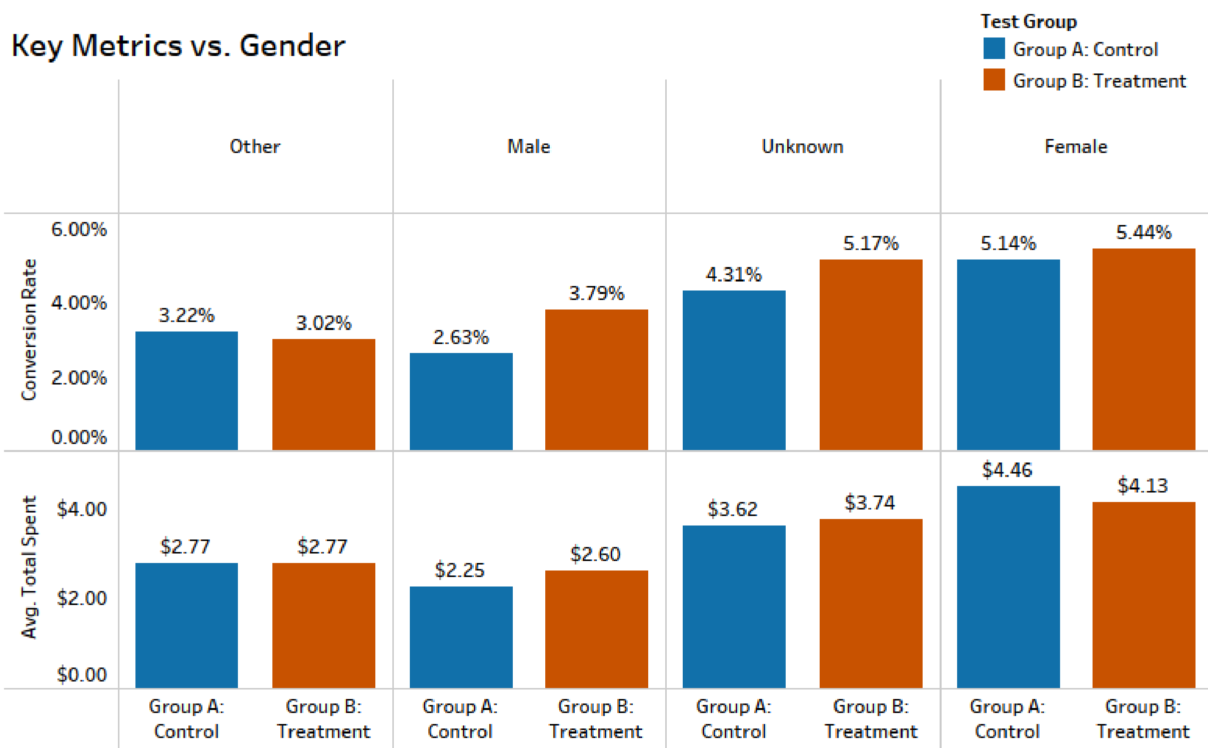
For the substantial conversion rate enhancement observed among Android users, the Growth team should sustainably elevate banner content quality. To achieve this, a deep dive into the elements that resonated with Android users can be achieved through surveys or feedback sessions. Insights garnered can then inform content enhancement, culminating in a more compelling call-to-action and heightened conversion rates.

## Analysis by Gender:

### Observations:

In the control group, it's evident that males exhibited the lowest conversion rate and average spending compared to all genders. However, a shift to the treatment group brings forth intriguing findings. Among male users in the treatment group, there was a striking 46.97% surge in the conversion rate and 17.69% in total spend.

On the other side, the impact on female users was less pronounced. Although females already demonstrated a relatively higher conversion rate in the control group, the treatment group registered only a marginal 5.60% increase.




### Suggestions:

The banner's impact is evident in a robust 46.97% conversion rate increase and 17.69% higher total spending among male users, emphasizing its alignment with this segment. Tailored marketing to males in the food category is recommended.

Conversely, female users experienced a -7.53% spending drop but a 5.60% conversion increase. To boost engagement and spending, deeper insights through surveys or feedback are advised.

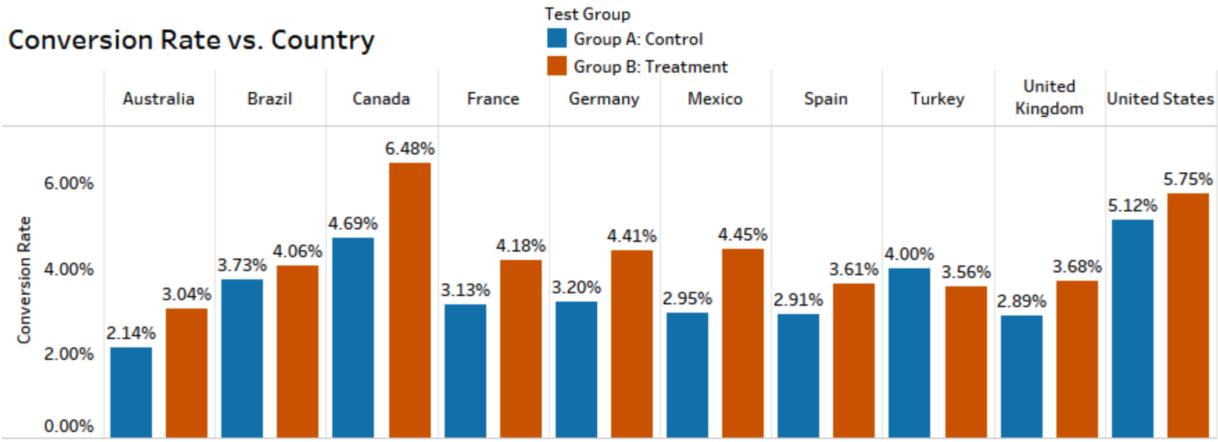
For the "Other" and “Unknown” group, deeper analysis is needed to uncover influencing factors. Strategies enhancing their food category engagement are crucial.

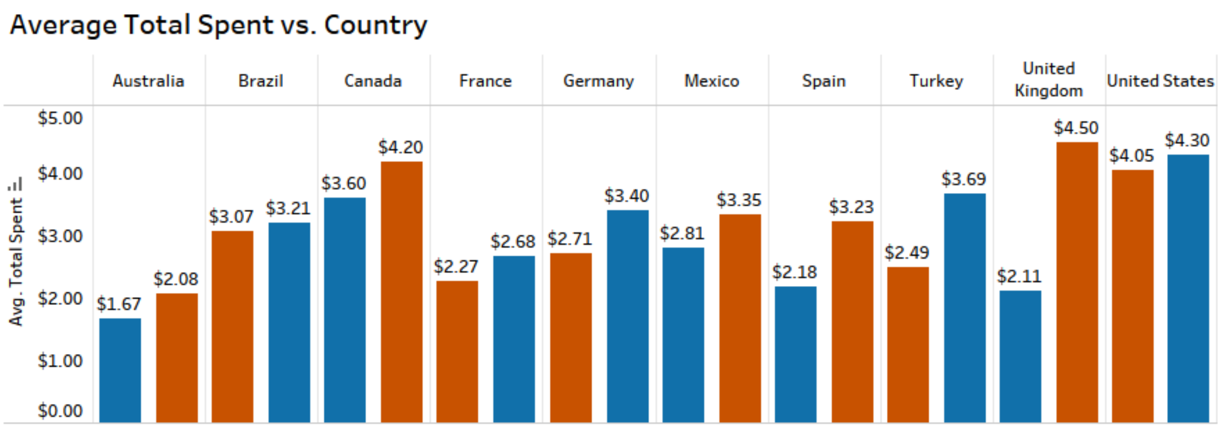
## Analysis by Country:

### Observations:

Diverse Conversion Rate Surges and Total Spending Growth:In the realm of conversion rates, Mexico and Canada have secured notable gains, with impressive boosts of 56.63% and 44.44%, respectively. These enhancements underscore a robust positive impact on user engagement and banner interaction, reflecting the efficacy of strategic targeting and content optimization.

Additionally, the United Kingdom has demonstrated an exceptional surge of 119.02% in total spending, while Spain has shown commendable growth of 48.32% in total spending. These figures collectively showcase the diverse yet impactful outcomes of our strategies across various regions.




### Suggestion:

Make the Most of Mexico and Canada's Progress: Leverage the notable conversion rate improvements in Mexico and Canada. Tailor marketing strategies and offers to match their preferences, driving engagement and revenue.

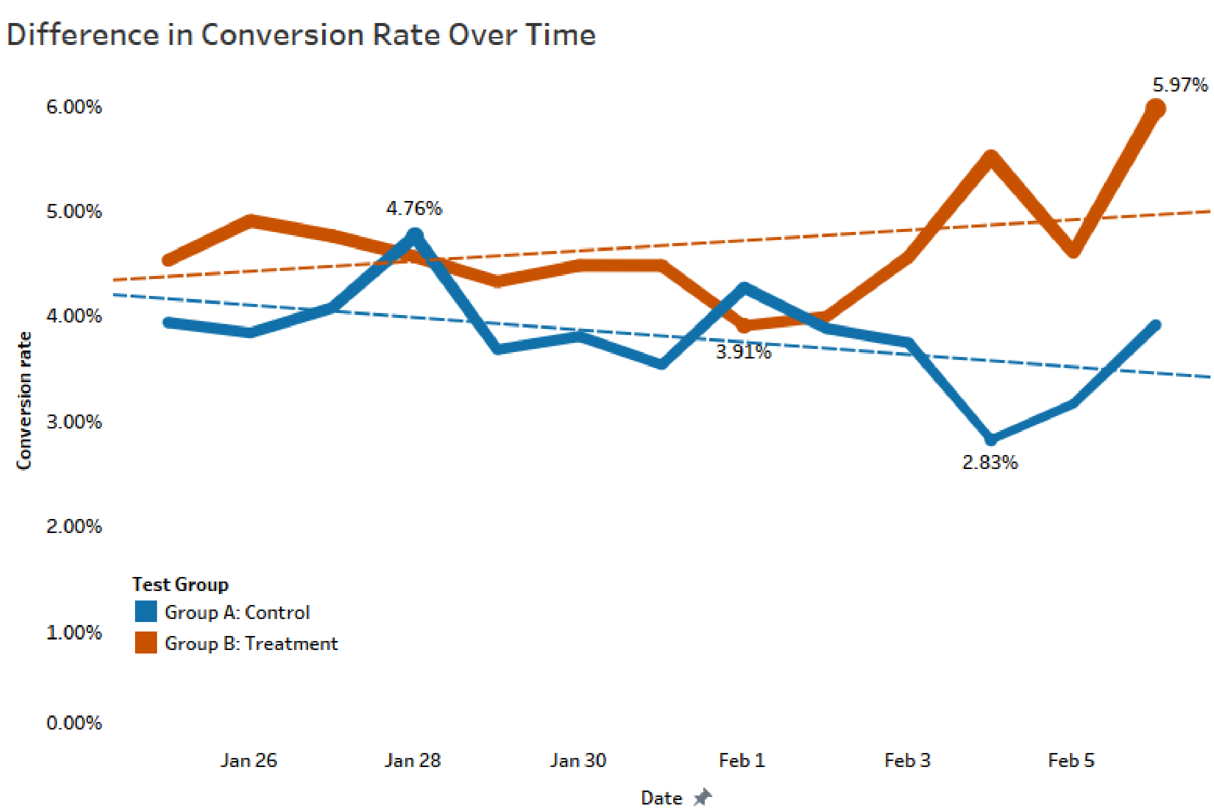
Understanding Turkey's Complex Spending Behavior: Investigate the factors behind Turkey's spending shifts, including a -9.46% decrease in conversion rates and a -31.22% reduction in total spending. Use comprehensive research and user feedback to identify opportunities for enhancing spending patterns and adapting strategies.

## Exploring for Novelty Effect:

Analyzing the likelihood of a novelty effect within the experiment involves a comprehensive assessment of the provided data.

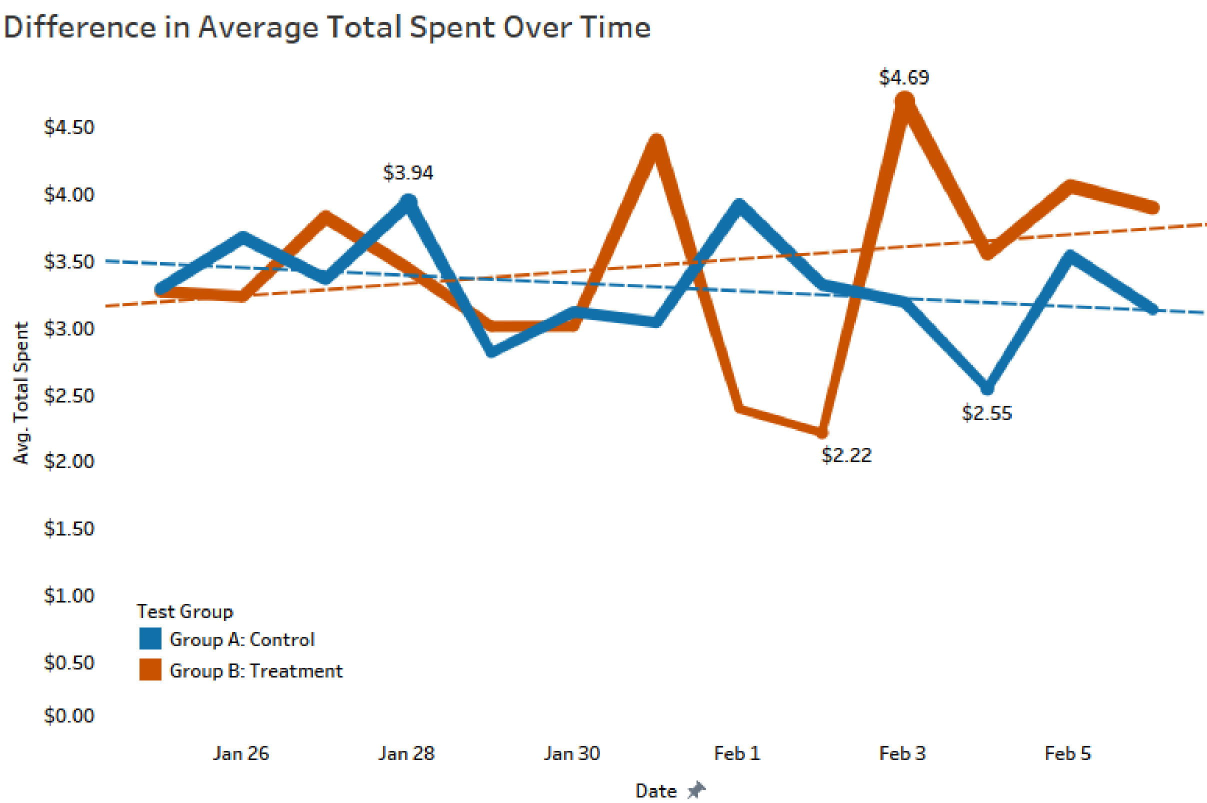
### 1. Conversion Rates Across Groups:

Now, delving into the conversion rates among all users in both groups over time presents a similar scenario. Comparable to the average spending patterns, the conversion rates lack consistent, notable disparities between the Control and Treatment groups. The minor oscillations observed do not indicate any novelty effect impacting conversion rates.



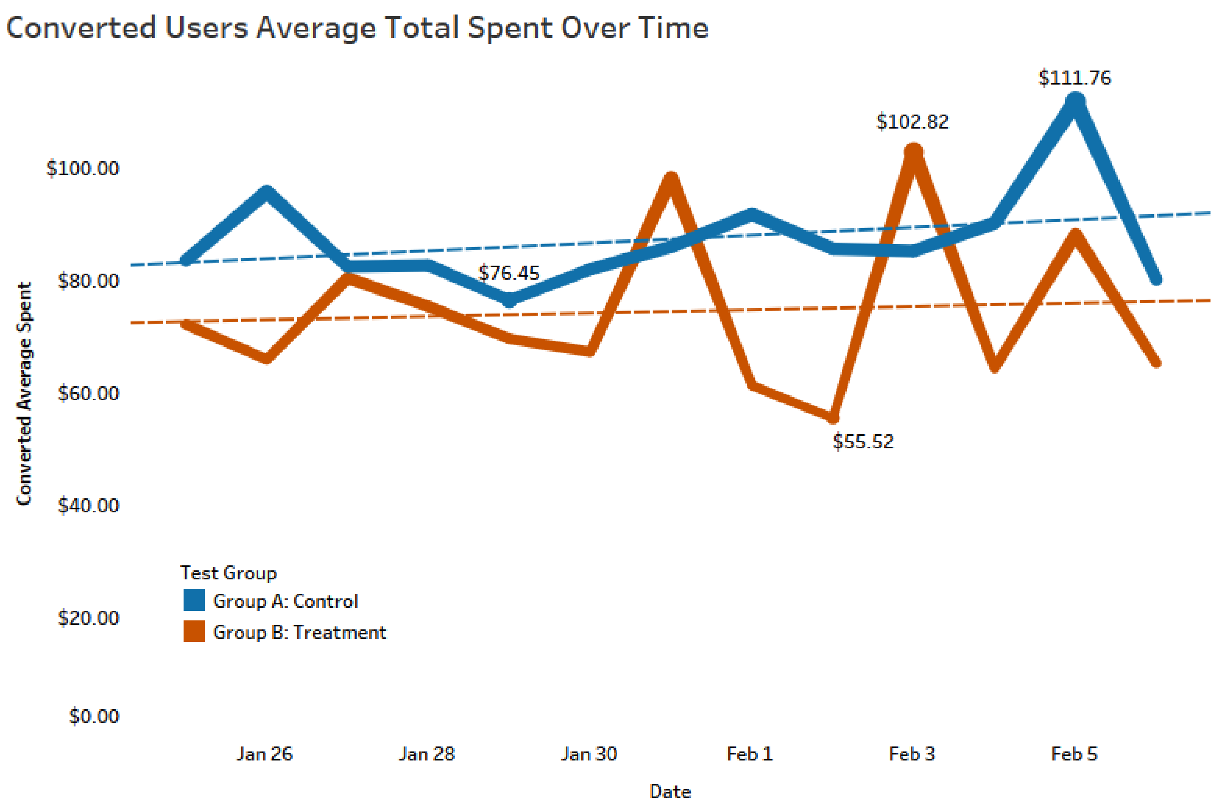
## 2. Average Spending Patterns Across Groups:

Examining the average amount spent by all users in both groups exposes certain fluctuations. However, it's vital to view these variations in the context of broader trends and not fixate on individual days. Visualizing this data illustrates that no consistent and prominent trends hint at a significant novelty effect.



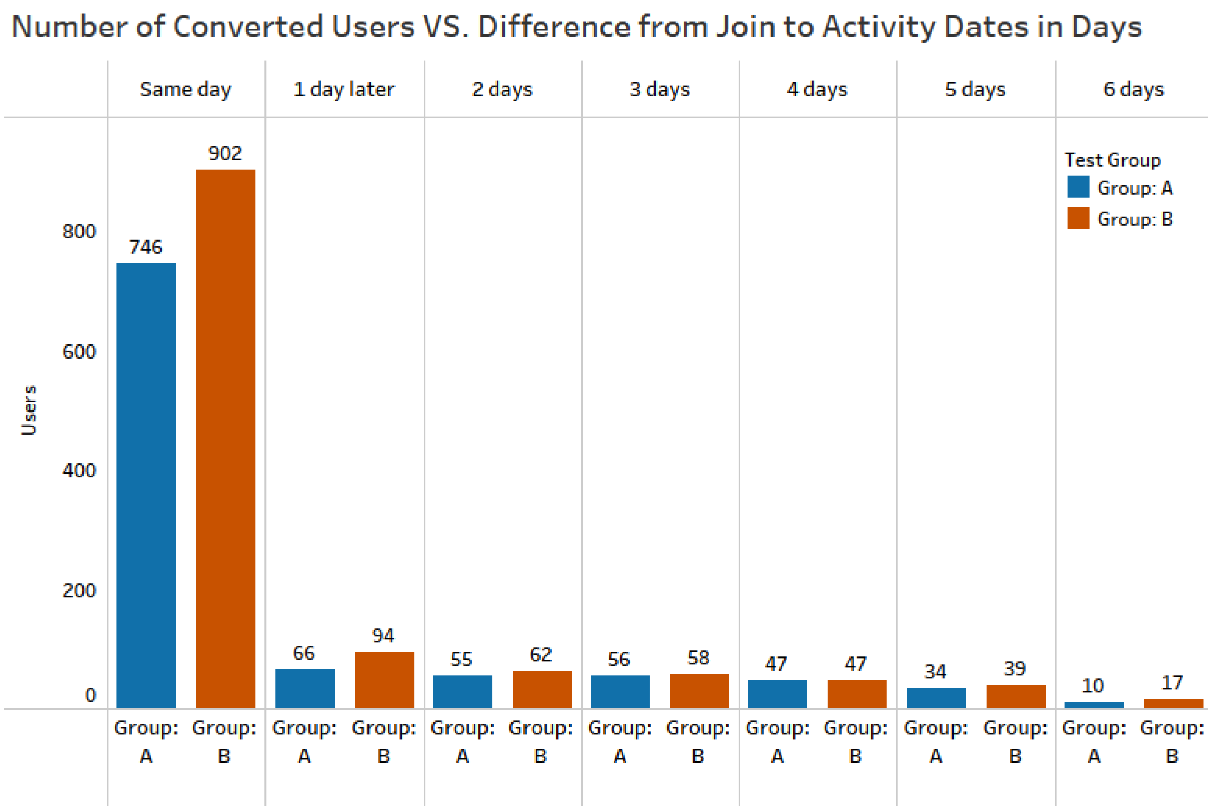
### 3. Average Spending Among Converted Users:

Examining the average spending exclusively among the users who have converted in both the Control and Treatment groups brings to light variations in patterns. Nevertheless, it's worth noting that there isn't a clear presence of a striking novelty effect in these observations. While there are indeed disparities between the spending behavior of the two groups, these differences, as they stand, don't demonstrate a uniform level of statistical significance.



## 4. Timing of User Purchases:

Finally, analyzing the time users from both groups take to make purchases following their participation offers insight. However, no substantial distinctions in the time to purchase between the Control and Treatment groups emerge. While some users make immediate purchases and others take a few days, these patterns do not suggest a novelty effect.

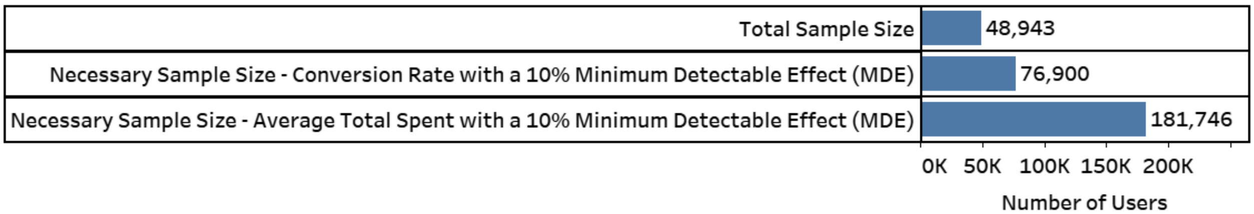



Upon meticulous analysis of the provided data, it becomes evident that the A/B test does not exhibit signs of a novelty effect. The metrics examined – average spending, conversion rates, spending among converted users, and purchase timing – do not consistently showcase noteworthy differences between the Control and Treatment groups.

As we assess these observations, it is crucial to explore how users might respond differently when exposed to new treatments, which is known as a novelty effect. By scrutinizing key metric differences between groups over time, we can identify any transient novelty effect. Such an effect, if observed, might imply that the banner's effectiveness is short-lived, prompting reconsideration of its launch viability.

*After conducting a thorough analysis of the A/B test data, it is evident that there is no substantial evidence indicating the presence of a novelty effect.*

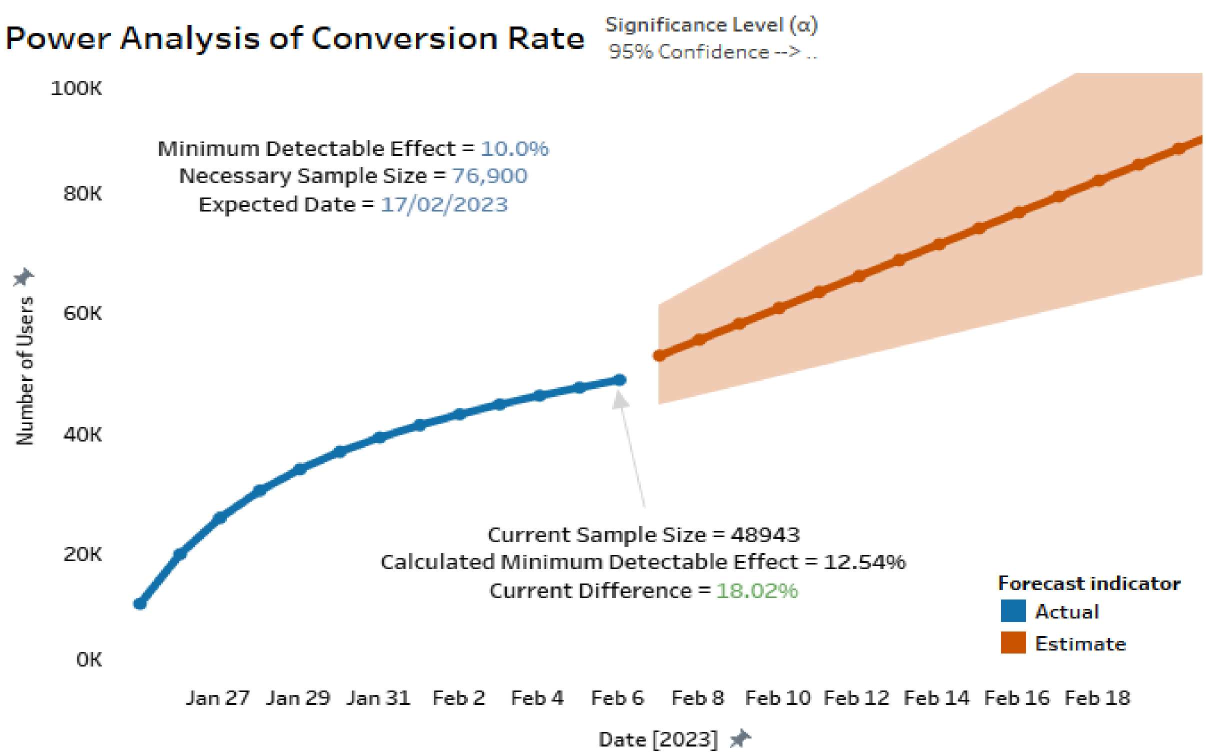
## Power Analysis:




The investigation centered on comparing conversion rates between treatment and control groups, emphasizing practical significance. Employing an online power analysis tool (Appendix link provided) and a two-sided 5% significance level, a sample size of 76,900 was recommended. This choice ensures the study's statistical power to detect the desired effect size reliably. The thorough analysis bolsters the study's credibility, confirming the appropriateness of the selected sample size. Impressively, the observed effect size in conversion rates significantly exceeded the 10% minimum detectable effect by a substantial margin of 18.02%.

The power analysis not only affirms the adequacy of the chosen sample size but also substantiates its capacity to accurately identify and validate the observed effect. This meticulous approach to power analysis and sample size determination fortifies the validity and practical implications of our findings, facilitating well-informed decisions for optimizing conversion rates.

Moreover, the calculated expected date of February 17, 2023, provides a valuable perspective on the experiment's timeline, offering insights into how much longer the experiment would need to run to reach the desired sample size. This projection aids in forecasting and resource planning, enhancing the experiment's overall efficiency and effectiveness.

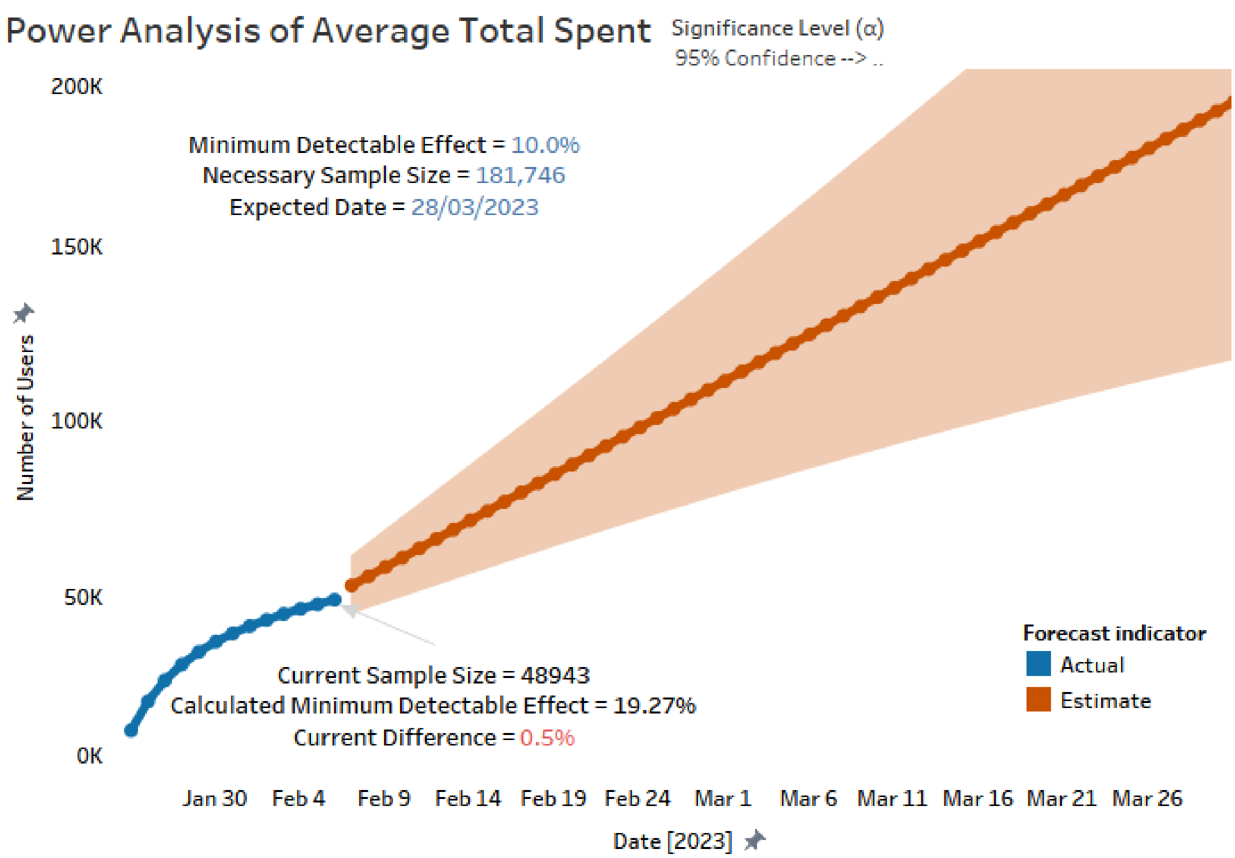


Additionally, a separate analysis was undertaken to compare two distinct means for the total amount spent. Here, the disparity between the means was established at 10% of the control group's mean, yielding a value of 0.337. The presumed standard deviation utilized for this analysis was 25.67 units.

Based on these parameters, the power analysis determined that a total sample size of 181,746 would be required. This sample size was calculated to achieve an 80% probability of detecting a genuine distinction in means of 0.337 units, while maintaining a 5% significance level (two-sided). The calculation considered the utilization of a combined standard deviation of 25.67 units (Calculator link is provided in the appendix).

However, in the context of average total spending, the effect size was notably reduced (0.5%), falling well below the minimum detectable effect of 10%. The power analysis revealed that capturing this subtle effect with the desired statistical potency would necessitate a considerably larger sample size. Nonetheless, even with a larger sample, the disparity in average spending between the control and treatment groups might continue to exhibit marginal significance.

Considering the projected expected date of March 28, 2023, strategic planning should focus on optimizing conversion rates, as the study demonstrates the statistical robustness of the chosen sample size for detecting substantial shifts in this metric.




The power analysis highlights that our sample size is suitable for comparing conversion rates with confidence. However, in terms of average spending, the small effect size and difficulty in obtaining a larger sample raise doubts about its practical impact. Importantly, focusing on conversion rates is key, as our data supports meaningful insights in this area. While average spending matters, its importance is limited due to small effects and challenges in getting a bigger sample. Therefore, our priority should be on strategies to enhance conversion rates and overall user engagement for revenue growth.

*In summary, the power analysis validates our sample size for detecting significant shifts in conversion rates. Yet, this statistical strength doesn't apply to average spending due to small effects and logistical constraints. Thus, our focus should shift towards boosting conversion rates.*

# Recommendation

Drawing upon an in-depth analysis of the A/B test outcomes along with supplementary insights, we possess a substantial foundation of information and statistical substantiation, enabling us to confidently advocate for the launch of the food and drink category banner. The A/B test findings manifest a noteworthy divergence in conversion rates between the control group and the treatment group, indicating a positive influence of the banner on driving conversions.

Nevertheless, it's imperative to acknowledge the intriguing trend observed in the average expenditure per user. While the treatment group exhibits a significantly superior conversion rate, the data hints at an interesting nuance – the surge in users from this group doesn't yield an equally proportionate revenue increase. This implies a potential shift in user behavior within the food and drink category, perhaps indicating a preference for lower-priced items.

The strategic choice to prioritize user acquisition over revenue per user comes with rationale. By attracting a broader audience through the food and drink banner, the potential for cultivating a larger customer base and fortifying brand interaction becomes palpable. Such a trajectory could potentially translate into heightened customer loyalty and recurrent purchases over the long haul.

Furthermore, the banner's launch carries a notable advantage in terms of cost-effectiveness. The minimal financial commitment required mitigates the financial risk linked to this decision. Even if the immediate impact on revenue per user might not be staggering, the budget-friendly nature renders it a feasible avenue to explore. This is particularly significant when one contemplates the potential long-range advantages arising from amplified user acquisition and heightened customer engagement.

However, it's also imperative to acknowledge the crucial aspect of optimizing revenue generation. To strike a harmonious balance, diligent monitoring of the banner's performance is warranted, complemented by comprehensive analyses delving into the spending behavior of users in the treatment group. This could involve segmenting the data based on various facets such as demographics, product preferences, and order values.

Propelled by these considerations, **I recommend launching the food and drink category banner**. Concurrently, it is advisable to undertake a more granular exploration of user spending patterns. This dual-pronged approach will furnish us with a deeper comprehension of the banner's effect on diverse customer segments, allowing for the tailored refinement of marketing strategies.

*In conclusion, the comprehensive A/B test analysis report supports and substantiates the proposal to introduce the food and drink banner.*

# Appendix

## SQL Queries:

- Click here to access all the queries that have been written for this project to extract CSV files.

## Spreadsheets:

Some of these files have been extracted in CSV format using SQL queries, while others have been created using data from CSV.

- Click here for hypothesis test initial analysis.
- Click here for non-SQL-based confidence interval data.
- Created two data sets for Novelty effect one and two.
- Created two data sets for Power analysis one and two.

## Tableau:

- This is the link for Tableau Public. Click here to see the visualization created for this project.

## Online Calculators:

- Click here for Sample Size Calculator for Conversion Rate.
- Click here for Sample Size Calculator for Comparing Two Independent Means.