# Evaluating the Impact of a New Banner on User Engagement: An A/B Testing Analysis for GloBox's Food and Drink Offerings

*Author: Christopher H. Kroll, Date: November 17, 2023*

## 1 Summary

This study examines the effect of a new banner on GloBox's e-commerce platform, specifically targeting their food and drink section. Conducted through an A/B test on their mobile website, users were divided into control and test groups, with the latter experiencing the banner. Key performance metrics, Conversion Rate (CR) and Average Amount Spent (AAS), were analysed using z-tests and t-tests. The results indicated a substantial rise in CR for the group exposed to the banner, without adversely affecting AAS. Additionally, the study noted changes in buying patterns and varying responses based on demographics and regions. These insights endorse the banner's deployment and highlight the importance of ongoing monitoring and a prolonged evaluation period for a thorough understanding of its impact.

## 2 Context

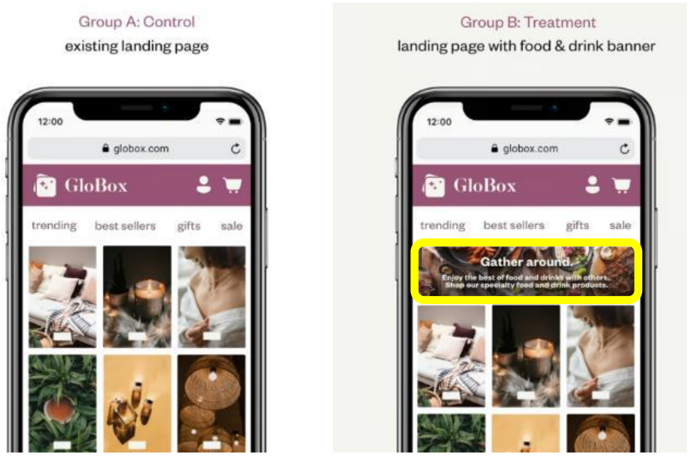
### 2.1 Evaluating Banner Impact on E-Commerce Metrics

GloBox is enhancing the e-commerce experience by analysing customer behaviour through an A/B test that assesses a new banner showcasing food and drink products. This test aims to determine the banner's influence on Conversion Rate and Average Amount Spent, to guide marketing strategies and make informed decisions on its implementation for revenue growth in this category.

### 2.2 A/B Test Setup

Below are the specifics of the A/B test:

- **Platform:** The experiment was exclusively conducted on the mobile website.
- **User Experience:** Upon visiting the GloBox main page, users were randomly assigned to either the control or test group. This marked their join date for the experiment. Depending on their group, the page either displayed the banner (test group) or remained as is (control group).
- **Conversion:** A user's decision to make a purchase, either immediately after joining the experiment or days later, was termed as a "conversion".




Images of the two website versions without (left) and with banner (right).

### 2.3 Dataset Overview

To conduct a thorough analysis of the A/B test's impact, we rely on a structured dataset that captures user demographics, their A/B test group assignments, and purchase activities. Here's an overview of the data tables and their respective columns:

**1. users (User Demographic Information):**

- **id:** Unique identifier for each user.
- **country:** ISO 3166 alpha-3 country code representing the user's location.
- **gender:** Specifies the gender of the user. Categorized as Male (M), Female (F), or Other (O).

**2. groups (A/B Test Group Assignment):**

- **uid:** Unique identifier corresponding to the user ID from the 'users' table.
- **group:** Specifies the test group the user has been assigned to.
- **join_dt**: Date on which the user was included in the test, i.e., the date they visited the page.
- **device:** The device type used by the user to access the page. Denoted as iOS (I) or Android (A)

**3. activity (User Purchase Activity):**

- **uid:** Unique identifier corresponding to the user ID from the 'users' table.
- **dt:** The date on which the purchase was made.
- **device:** Device type on which the purchase was conducted. Denoted as iOS (I) or Android (A).
- **spent:** Total purchase amount in USD for that day.


Image of the Entity Relationship Diagram of the GloBox Database.

**Key notes**

- Every user in the dataset has been assigned to one of the A/B test groups.
- It's important to note that not all users have made a purchase.
- The purchase activity captured encompasses all product categories, not exclusively the food and drink segment.

# 3 Methodology

To thoroughly investigate the effectiveness of the new banner on user behaviour, our team employed a comprehensive methodology, utilizing a variety of tools and statistical analysis techniques. Below is an overview of our approach:

## 3.1 Tools Used

- **Tableau:** Aided in visualizing the data and drawing insights from the user interactions and purchase behaviour.
- **SQL in Beekeeper:** Facilitated querying and aggregating the data, ensuring we could effectively analyse user behaviour and purchase patterns.
- **Spreadsheets:** Used for additional data manipulation and analysis, providing a flexible environment for calculations and comparisons.
- **ChatGPT Model 4:** Assisted in SQL coding, data interpretation, providing insights, and generating content for analysis and reports.

## 3.2 Statistical Analysis

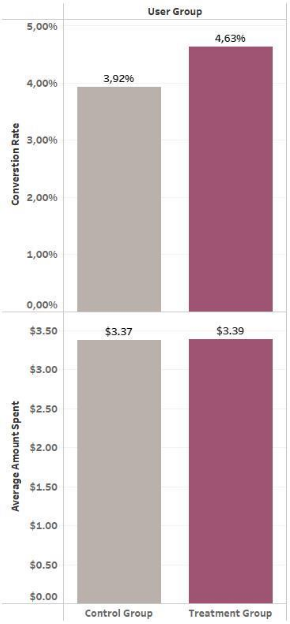
- **Conversion Rate (CR):** A z-test was performed to compare the Conversion Rates between the control and variant groups. This test was chosen for its suitability in comparing sample proportions. The analysis was conducted at a significance level of $\alpha$=0.05. The effect size was measured using Cohen's *h*, providing a quantifiable measure of the impact's magnitude.
- **Average Amount Spent (AAS):** At the same significance level ($\alpha$=0.05), a t-test was conducted to compare the Average Amount Spent between the two groups. Given the nature of the data, this test was appropriate for comparing the means of two independent samples. The effect size was quantified using Cohen's *d*, offering a numerical representation of the difference's magnitude.
- **Novelty Effect**: In our methodology, we thoroughly examined the possibility of a novelty effect following the banner's introduction. This involved an analysis of the initial conversion rate trends to identify any short-term spikes that could be attributed to novelty. Our approach was designed to distinguish between genuine, sustained changes in user behaviour and those potentially influenced by the initial novelty of the banner.
- **Power Analysis and Forecasting:** To determine the necessary duration of the experiment for capturing a Minimum Detectable Effect (MDE) on CR and AAS, power analysis and forecasting were conducted. These analyses guided our recommendations on the experiment's duration to ensure robust and reliable results.

## 3.3 Rationale Behind Decisions

The chosen tools and statistical methods were driven by the need for a comprehensive understanding of the banner's impact, ensuring accuracy and reliability in our findings. The power analysis and forecasting were crucial in guiding the experiment's duration, while the use of statistical tests ensured we could confidently assess the banner's impact on key metrics.

# 4 Results

Our A/B test on the introduction of the new banner brought forth enlightening insights into user engagement and spending patterns. Notably, at a significance level of $\alpha = 0.05$, the Conversion Rate (CR) witnessed a significant enhancement with the banner's implementation. The control group posted a CR of 3.92%, but with the banner, the variant group reached an impressive 4.63%, resulting in an observed mean difference of 0.71%. The 95% Confidence Interval for this difference was [0.35%, 1.07%]. The z-test yielded a z-statistic of 3.8643 and a $p < 0.001$, indicating that the observed difference is statistically highly significant. The effect size, measured using Cohen's *h*, was 3.86, indicating that the observed effect is positive and very large. This suggests that the banner has a positive impact on conversions and could attract more or new customers, looking for food and drink offerings especially.




CR and AAS by Group

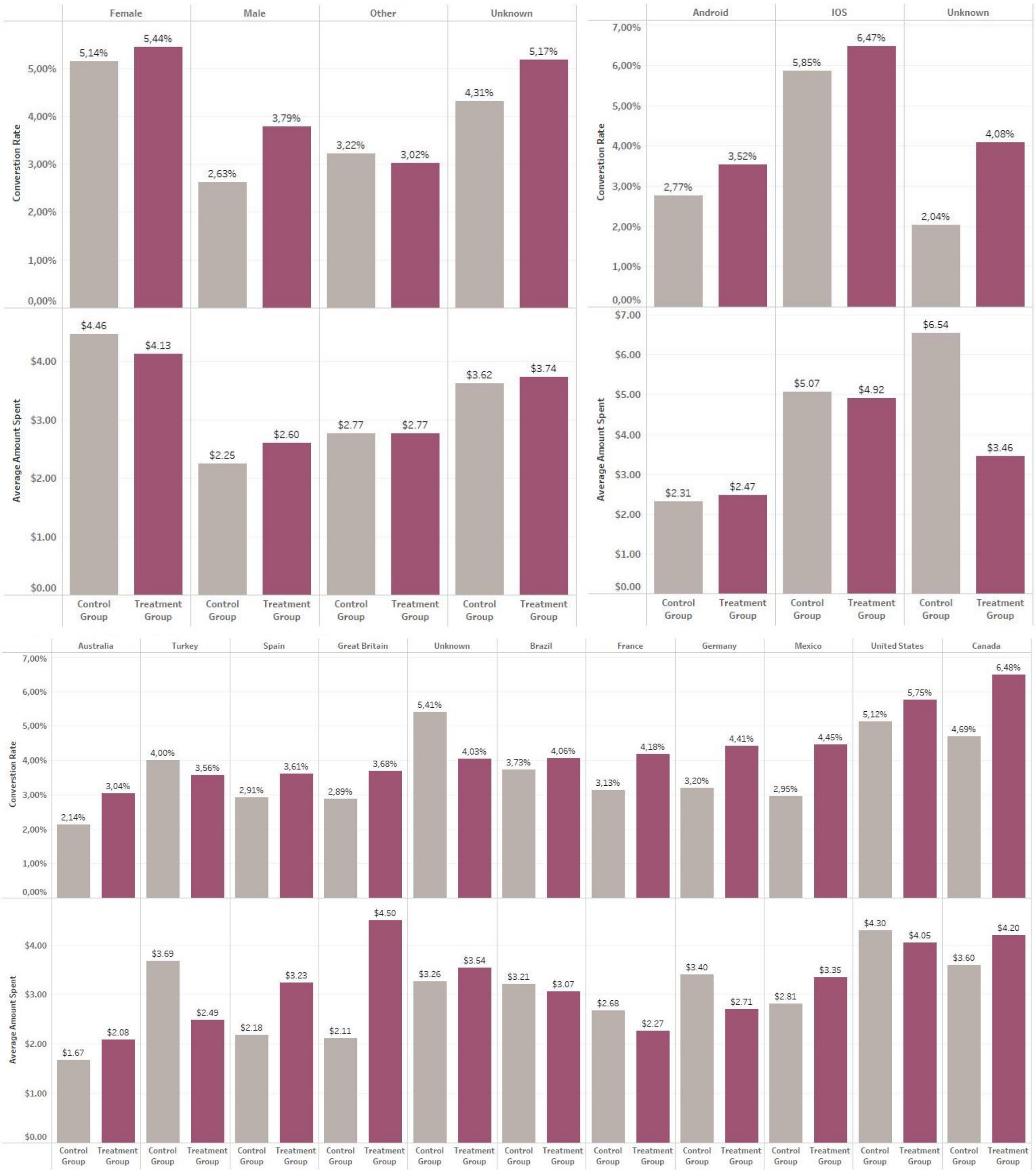

95% Confidence Interval of the Difference of CR


95% Confidence Interval of the Difference of AAS

Turning to the Average Amount Spent (AAS), the results were more subdued. At the same significance level (α = 0.05), the control group had an average spending of $3.37, while the variant group edged slightly higher at $3.39, resulting in a minimal observed mean difference of $0.02. The 95% CI for this difference was [$-0.44, $0.47]. The t-test produced a t-statistic of 0.0704 with 24599 degrees of freedom and a p-value of 0.943. Statistically, this difference was minimal and insignificant, as evidenced by the effect size (Cohen's *d*) of 0.0006, which was well below the threshold for even a small effect size. Importantly, there was no observed negative impact on the Average Amount Spent.

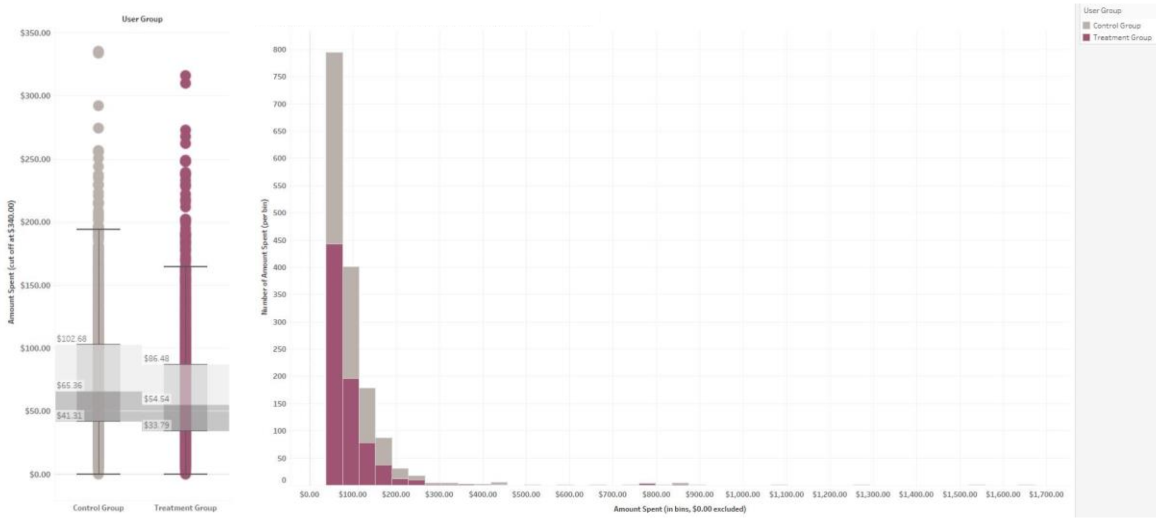
In GloBox's luxury-dominated context, introducing lower-priced food products might have impacted conversion rates and Average Amount Spent (AAS). The consistent AAS, despite a diverse price range, suggests that adding food items increases purchase frequency without affecting overall user spending. However, the steady AAS alongside a higher conversion rate implies a shift in the treatment group's buying patterns, possibly favouring these affordable food items over luxury goods, hinting at potential product cannibalization. Therefore, the static AAS, despite higher conversion rates, indicates that the actual spending per purchasing customer is likely lower in the treatment group, especially when non-purchases are excluded.

Delving deeper into the demographics, both males and females showed a favourable response in terms of CR. However, while male spending increased, female spending registered a decline. From a geographical standpoint, regions such as Australia, Spain, Great Britain, Mexico, and Canada responded positively across both metrics. Yet, Brazil, France, Germany, Turkey, and the United States painted a mixed picture, endorsing the banner in terms of CR but pulling back when it came to AAS.



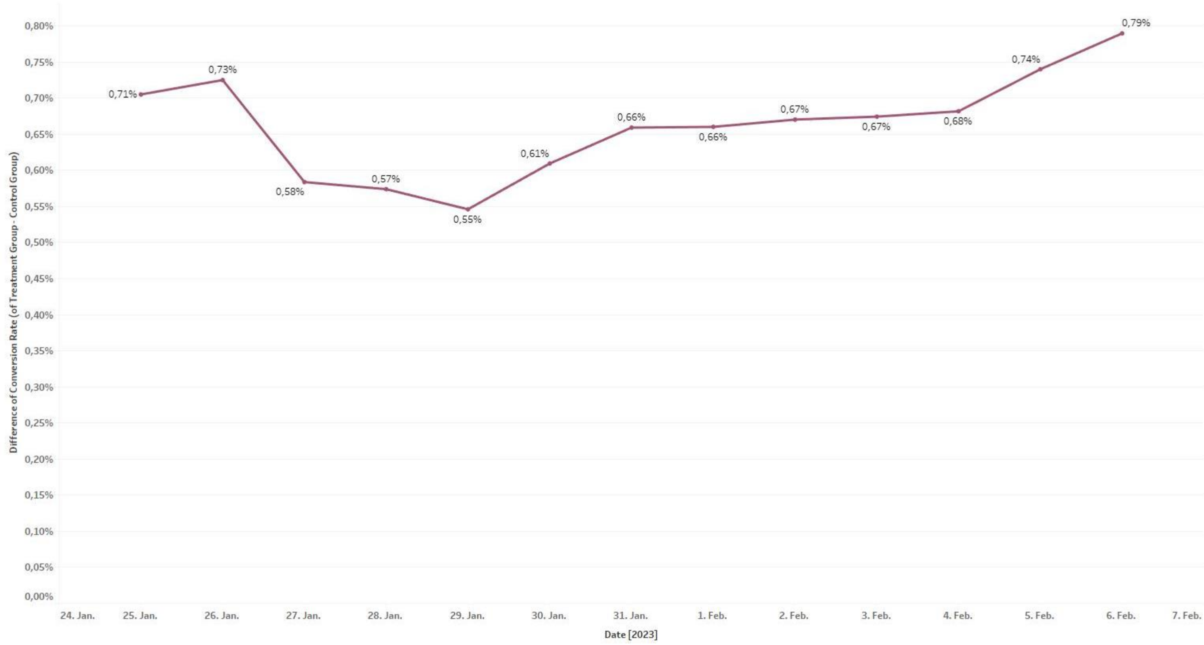

CR and AAS by Gender x Group, Device x Group and Country x Group

Our analysis of the purchase behaviour revealed a strong right-skewed distribution for AAS, signifying a tendency among users to spend smaller amounts. However, the treatment group displayed a mild contraction in this distribution, indicating a potential shift in their purchasing habits.



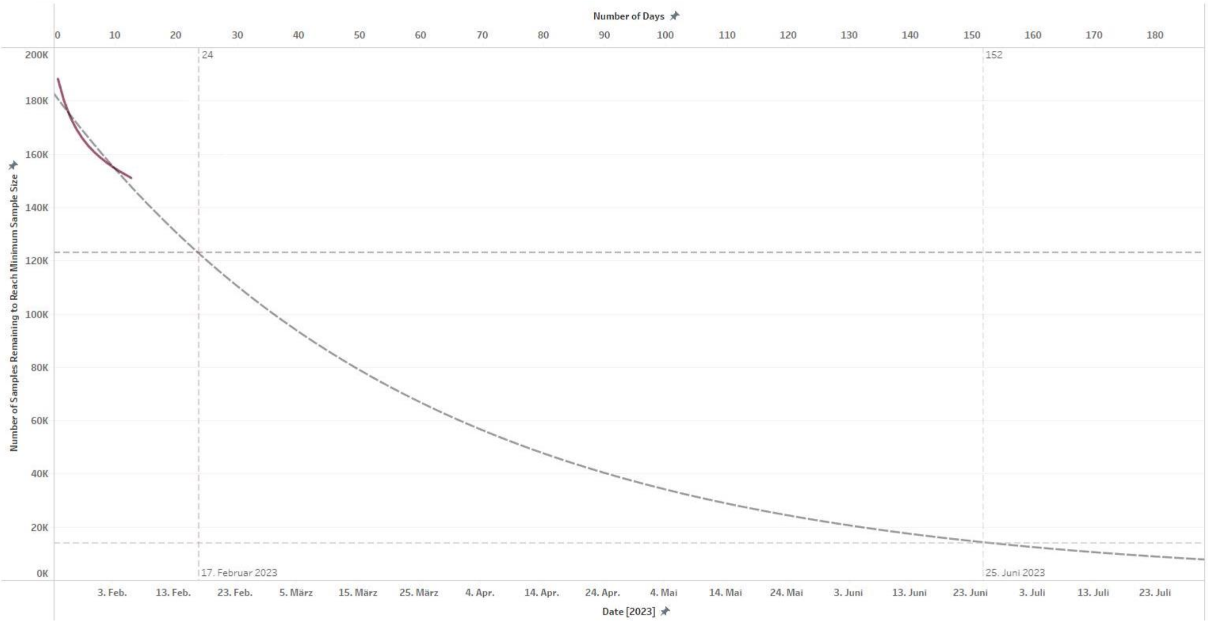

Distribution of Amount Spent per user by Group

In the early phase of the new banner's rollout, specifically the first two days, we observed a pronounced increase in CR, with a difference (Δ) of 0.73% in favour of the treatment group. This initial rise, more likely due to small sample variability, challenges the notion of a novelty effect. As the test continued, the conversion rate difference narrowed, further questioning the novelty effect theory. By day five, the CR began to stabilize, registering a difference (Δ) of 0.55%, more closely mirroring the control group. Interestingly, after this stabilization, the rate began to ascend again. By day 13, the difference soared to Δ=0.79, outpacing even the initial spike. This indicates that beyond the initial allure, the banner has a sustained and potentially increasing benefit over time.




Difference of Conversion Rate (of Treatment Group - Control Group)

Lastly, power analysis underscored the importance of a lengthier observation period. To fully grasp the banner's ramifications and ensure a Minimum Detectable Effect (MDE) of 10%, we would need at least 24 days for CR and a heftier 152 days for AAS.




Forecast on Experiment Duration necessary to reach Minimum Sample Sizes
(for MDE ≥ 10% of Conversion Rates and Average Amount Spent)

# 5 Recommendations and Conclusions

Following an exhaustive analysis of the A/B test results associated with the new banner's introduction, our team presents this report detailing our recommendations. This report is structured to offer insights derived from the data, the broader business implications, and to provide clear guidance for the next steps.

## 5.1 Launch the Banner

The data advocates for the banner's introduction, given the significant increase in the Conversion Rate and the absence of any adverse impact on the Average Amount Spent. Furthermore, business considerations, spanning operational costs to potential revenue growth, underscore this recommendation. It is anticipated that the banner, spotlighting the company's burgeoning food and drink offerings, will not only retain the existing customer base but also entice new users.

## 5.2 Ensuring Continuous Monitoring and Analysis

While the banner's initial reception has been positive, it's imperative to maintain an ongoing pulse check on its performance. Detailed analyses targeting specific demographics—gender, device type, and geographic location—will deepen our understanding of its impact. Moreover, studying patterns in the Average Amount Spent distribution will offer insights into user spending behaviour, essential for refining future strategies.

### 5.3 Commitment to an Extended Observation Period

Our power analysis has pinpointed the necessity of a longer observation window to capture a comprehensive understanding of the banner's impact. This entails:

- A 24-day observation for the Conversion Rate.
- An extended 152-day period for the Average Amount Spent.

Such an approach ensures that we glean the Minimum Detectable Effect (MDE) of at least 10%, thereby solidifying our conclusions.

### 5.4 Preparedness for Future Challenges

In the fluid landscape of e-commerce, user behaviours and preferences are ever-evolving. As such, we must remain vigilant, ready to recalibrate our approach should future data suggest any unforeseen negative implications. A specific area of focus would be the Average Amount Spent, ensuring that the banner continues to bolster, not hinder, revenue streams.

### 5.5 Conclusion

In summation, the introduction of the new banner, underpinned by robust data and strategic business considerations, presents an exciting avenue for GloBox to enhance its user experience and market reach. However, as with all strategies, it mandates ongoing scrutiny and the agility to adapt based on fresh insights and user feedback.

## 6 Appendix

### 6.1 Weblinks

- Basic Dataset and Working Spreadsheets: https://docs.google.com/spreadsheets/d/1G9I47rmIcQc1Yk_4BRPd9fC7Lqa_73kN8bowONlrCcA
- Novelty Effect Dataset: https://drive.google.com/file/d/1PplC8Zw6_NTsq89KseltrCqYjjKM6mKK
- Tableau workbook: https://public.tableau.com/app/profile/christopher.kroll/viz/MasteryProjectI/5_4_b_Cohens_d_AAS

## 6.2 SQL Queries

### 6.2.1 SQL Query returning the Basic Dataset

```
SELECT
    u.id,
    u.country,
    u.gender,
    g.device AS device_visit,
    g.group AS user_group,
    COALESCE(SUM(a.spent), 0) AS sum_spent,
    CASE
        WHEN COALESCE(SUM(a.spent), 0) > 0 THEN 1
        ELSE 0
        END AS is_converted
FROM users AS u
LEFT JOIN groups AS g
ON u.id = g.uid
LEFT JOIN activity AS a
ON g.uid = a.uid
GROUP BY u.id, u.country, u.gender, g.device, g.group;
```

### 6.2.2 SQL Query returning the date and the metrics for each group in separate columns

```
-- Step 1: Creating join_dt_agg
WITH join_dt_agg AS (
    SELECT
        g.join_dt,
        g.group AS test_group,
        COUNT(DISTINCT g.uid) AS user_count
    FROM groups g
    WHERE g.group IN ('A', 'B')
    GROUP BY g.join_dt, g.group
    ),

-- Step 2: Creating convert_dt_agg
convert_dt_agg AS (
    SELECT
        a.dt,
        g.group AS test_group,
        COUNT(DISTINCT a.uid) AS converted_user_count
    FROM groups g
    JOIN activity a ON g.uid = a.uid AND a.spent > 0
    WHERE g.group IN ('A', 'B')
    GROUP BY a.dt, g.group
    ),
```

```sql
-- Step 3: Creating cumulative_users
cumulative_users AS (
    SELECT
        COALESCE(a.dt, j.join_dt) AS dt,
        COALESCE(a.test_group, j.test_group) AS test_group,
        COALESCE(SUM(j.user_count) OVER (PARTITION BY j.test_group
        ORDER BY j.join_dt), 0) AS cum_users,
        COALESCE(SUM(a.converted_user_count) OVER (PARTITION BY
        a.test_group ORDER BY a.dt), 0) AS cum_converted_users
    FROM join_dt_agg j
    FULL JOIN convert_dt_agg a ON j.join_dt = a.dt AND j.test_group =
    a.test_group
    ),

-- Step 4: Creating cumulative_conversion
cumulative_conversion AS (
    SELECT
        dt,
        test_group,
        CASE
            WHEN cum_users = 0 THEN 0
            ELSE cum_converted_users * 1.0 / cum_users
            END AS cum_conversion_rate
    FROM cumulative_users
    )

-- Step 5: Final Result
SELECT
    a.dt,
    a.cum_conversion_rate AS conversion_rate_A,
    b.cum_conversion_rate AS conversion_rate_B,
    b.cum_conversion_rate - a.cum_conversion_rate AS cum_diff
FROM cumulative_conversion a
JOIN cumulative_conversion b ON a.dt = b.dt
WHERE a.test_group = 'A' AND b.test_group = 'B'
ORDER BY a.dt;
```