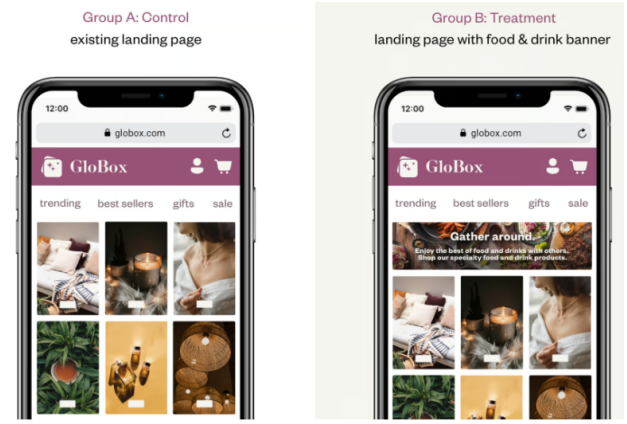
GloBox A/B Test Report

By: Oyieyo Silas **Date: 22/12/2023**

Updated Summary

The task aimed to analyze the GloBox A/B test result between the control and treatment. Based on the Conversion rate's hypothesis test on the estimators, I recommend launching the experiment.



PROJECT CONTEXT

"GloBox operates as an online market platform, specializing in curating and selling unique, high-quality products globally. The company's primary focus revolves around boutique fashion items and high-end decor products, having established a strong reputation among its customer base. However, GloBox aims to increase awareness of its rapidly expanding food and beverage offerings.

To enhance its homepage and facilitate customer engagement, GloBox is considering the addition of a dedicated product button. The company launched an A/B test experiment to test the influence. This experimental technique is commonly utilized in businesses to compare the performance of existing and new products. The groups consist of the control group (A), which does not view the banner, and the exposed group to the new feature (B).

The data required for analysis was extracted from the GloBox SQL Database. The dataset spans from January 25th, 2023, to February 6th, 2023."

HYPOTHESIS

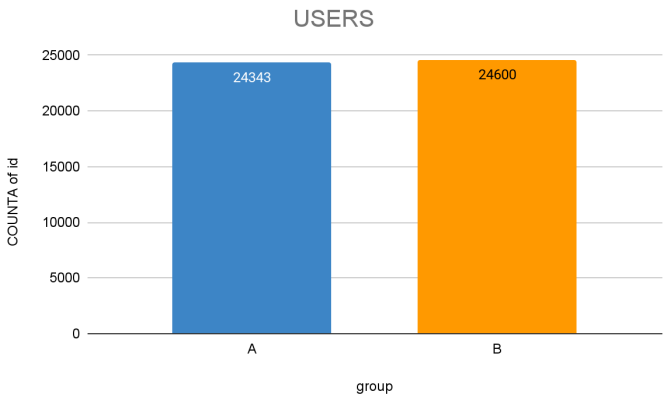
H0; The conversion rate of (Ac) equals the conversion rate of (Bc), (H0:Ac=Bc)

H1.The conversion rate of (Ac)A is not equal to the conversion rate (Bc),(H1: Ac!=Bc)

RESULTS

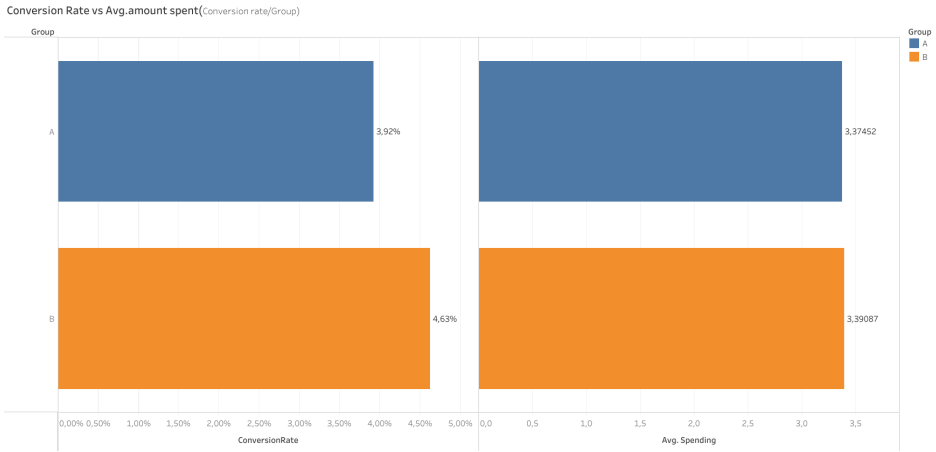
A. Proportions of the overall population.

Based on the test period, data contained 48943 distinct users, 24600 in the control and 24343 in the test group.



B. Conversion Rate and Spending Rate Performance.

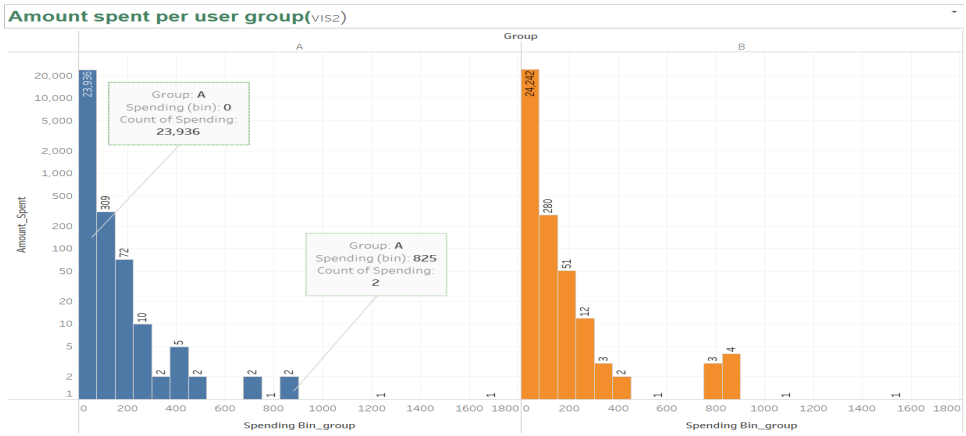
Despite having relatively lower spending than control(A), Treatment(B) demonstrates notably superior performance in both estimators.



C. The distribution of spending per user group

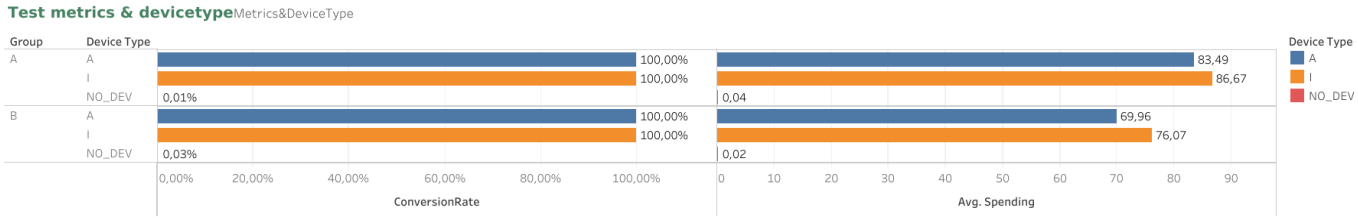
A decrease in spending is observed with an increase in the spending bin_group using a bin size of 75.

Both estimators exhibit a resemblance. NOTE: the graph is on a logical scale.



D. Comparing the estimators against the device

The conversion rate for users with Android (A) or iOS (I) devices stands at only 0.03%, whereas users with an Unknown device (No_DEV) do not contribute to the conversion.

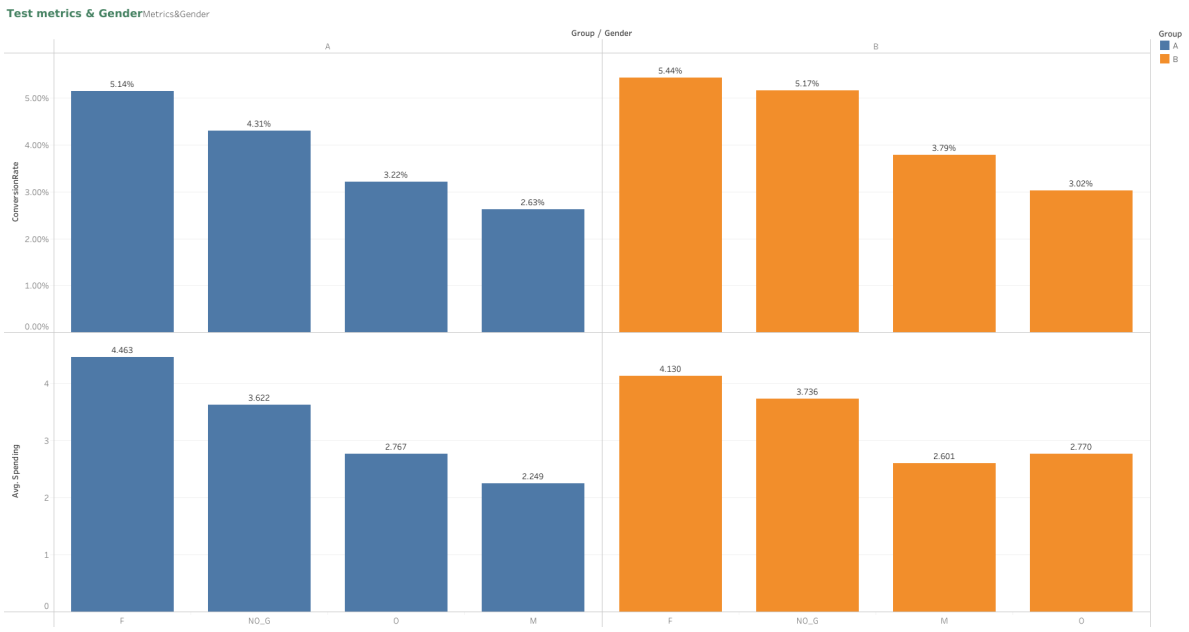


E. Comparing estimators against gender

Females exhibit a notably higher conversion rate in both groups, while the 'O' gender presents the lowest conversion rate.

Across gender groups, the test group demonstrates a higher conversion rate in nearly all categories.

An assumption can be made that women tend to prefer fashion and engage in online shopping more frequently.

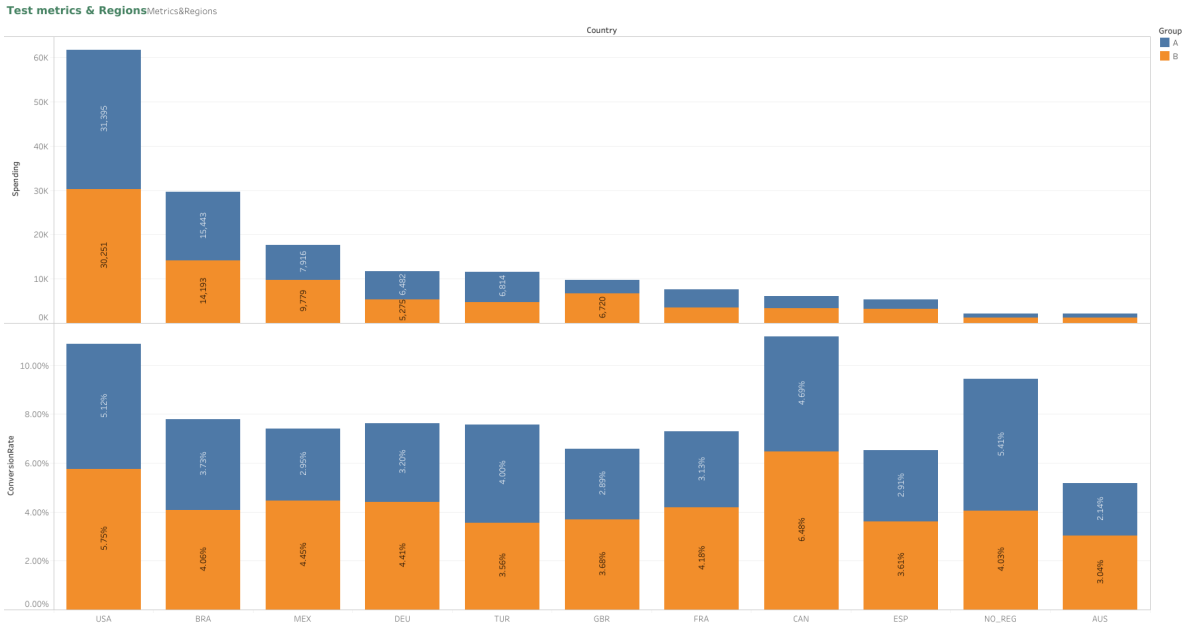


F. Relationship between user's country and estimators

The USA exhibits the highest spending among the groups and ranks second in conversion rates, following Canada.

Canada demonstrates a conversion rate closely approaching the USA's, yet it maintains relatively lower expenditure records.

Group B displayed slightly elevated conversion rates when compared to the control group.



G. Hypothesis Testing

- **95% Confidence Interval on Conversion rate per group.**

  The p_value of 0.000115 at 95% is less than Alpha 0.05. Therefore, we reject the null hypothesis, which states conversion rates between A and B are equal. The Test Statistic does not fall within the interval values between 0.0034831 and 0.01065657. We conclude that average conversion rates are different.

  Conversion rates have a significant mean difference (Sample Statistic) of 0.00707 between the test groups.

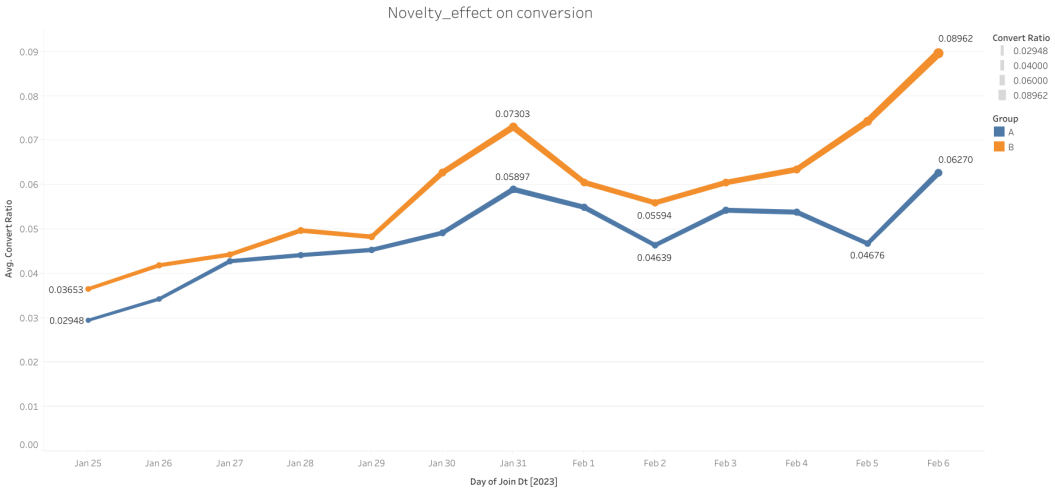
- **Test for 95% Confidence Interval on Spending rate per group.**

  The p_value of 0.943647 at 95% confidence is more than 0.05. Reject the alternative hypothesis that there is a difference between spending per group. We confirm the decision by the Test Statistics -0.0707, falling within the interval values between -0.4385189 and 0.4713189. The sampling statistic of 0.01635 between the means becomes insignificant.

H. Novelty Test

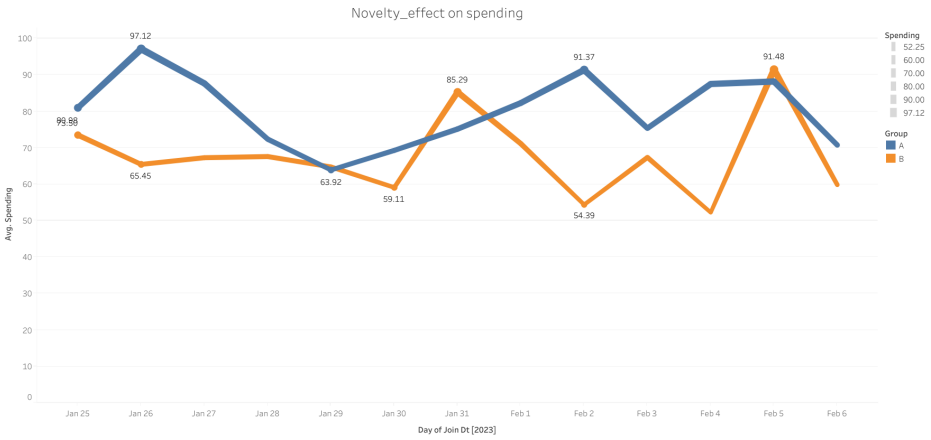
a)On Conversion rate

Initially, both groups had a relatively low conversion ratio, but they diverged around January 30 as group B recorded an increased rate. Test group (B) experiences a significant increase in the conversion ratio, peaking around January 31 before a decline in February 2. The control group (A) depicts the same projection but with a lower conversion rate. **B has a higher conversion rate than A. There is no indication of a novelty effect.**



b)On Spending rate

Both groups exhibit fluctuations in their spending throughout this period. Group A starts with higher spending but experiences a significant drop around Jan 28 before recovering and peaking on Feb 3. In contrast, Group B begins with lower expenditure, sees an increase around Jan 28, and peaks earlier than Group A on Feb 1. **The behavior in both groups may be due to other factors, not due to the new project. There is no indication of a novelty effect.**



Conclusion

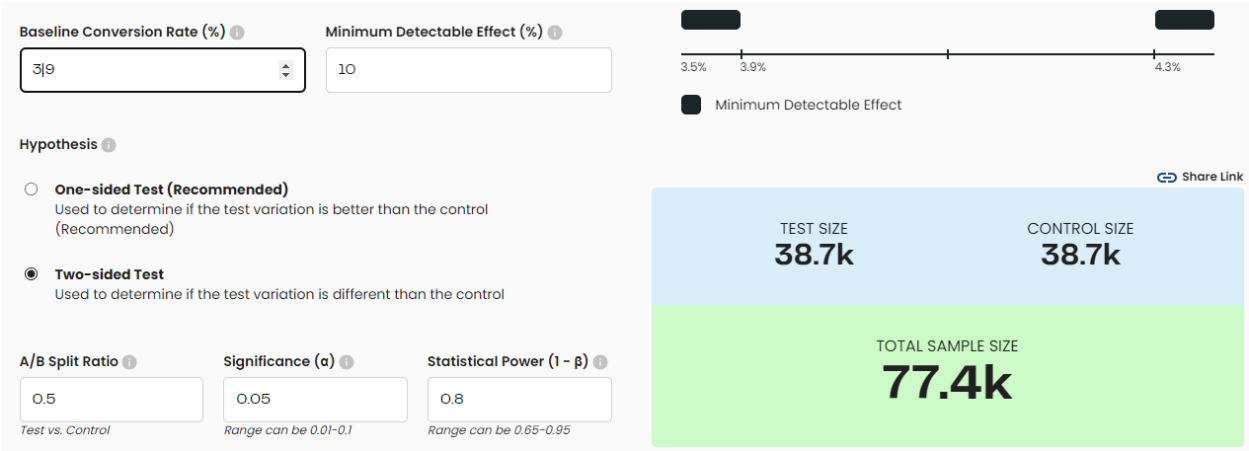
Based on the hypothesis test and the conclusion drawn from the sample statistic, I infer the decision concerning conversion rates. The substantial positive disparity in the average conversion rates between groups A and B presents sufficient evidence supporting the anticipation of a positive impact from

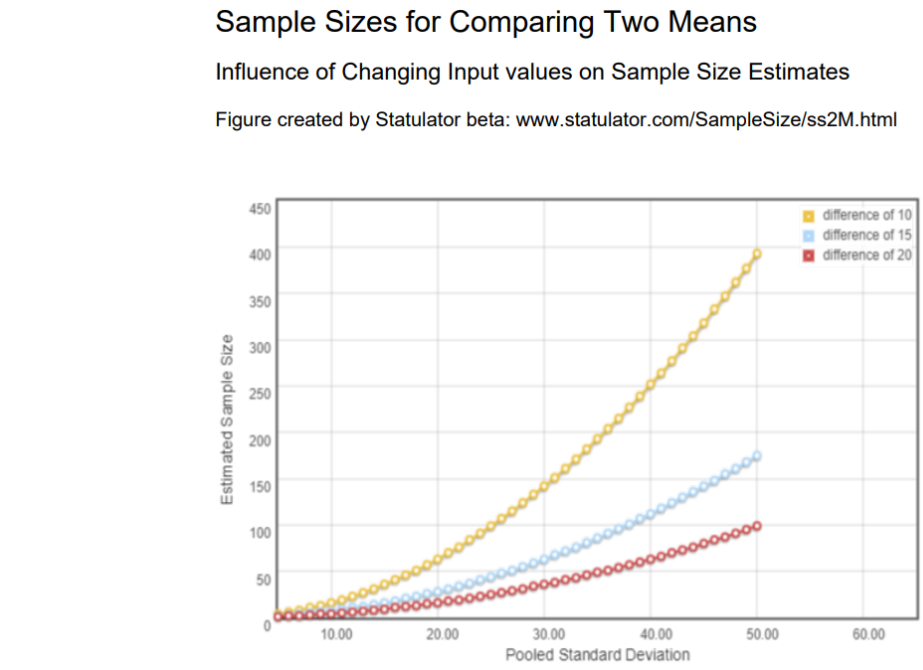
the new project. However, there is no significant variance in average spending between these groups, so external factors unrelated to the new project might influence the spending amount.

When examining conversion rates relative to gender, the treatment group (B) demonstrates a higher percentage across all genders than group A. A similar distribution is observed when comparing the average conversion rates across different regions.

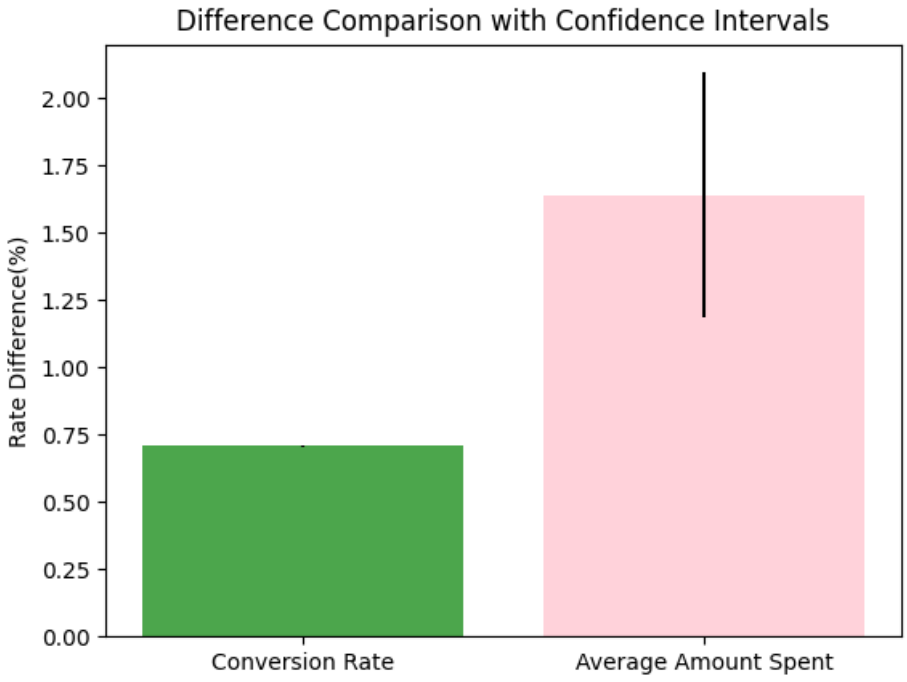
While the novelty effect lacks justification by considering both estimators, a noticeable daily conversion rate increase is evident compared to B. However, the subsequent fluctuation in the spending curve lacks predictability. In conclusion, there is an indication of the Novelty effect.

The control group exhibits a baseline conversion percentage of 3.9%. This rate serves as the benchmark for comparison. We assess whether the sample size is sufficient by employing a chosen minimum detectable effect of 10% and an 80% power. It is determined that over half of the analyzed sample size is necessary to achieve the desired 10% minimum noticeable effect. The sample size essential from a pooled standard error of 0.232 is less than the sample size used in the experiment on spending.





The confidence intervals between A and B in conversion and spending rates show a positive difference. In the treatment group, the estimator rates are higher than the control, with Spending showing a tremendous difference.



In conclusion, I recommend proceeding with the experiment's rollout.

Codes

1. **Downloaded Grouped CSV;**

```
SELECT u.id, COALESCE(u.country,'NO_REG') AS country, COALESCE(u.gender, 'NO_G') AS
gender,
       COALESCE(a.device,'NO_DEV') AS device_type, g.group,
       CASE WHEN a.spent > 0 THEN 'YES' ELSE 'No' END AS conversion,
       COALESCE(SUM(a.spent), 0) AS Spending
FROM users U
LEFT JOIN activity a
ON u.id = a.uid
LEFT JOIN groups g
ON g.uid = u.id
GROUP BY 1,4,5,6;
```

2. **Novelty Test code**

```
-- Step 1: Creating join_dt_agg
CREATE TEMPORARY TABLE join_dt_agg AS
SELECT g.join_dt, g.group, COUNT(DISTINCT u.id) AS user_count
FROM users u
INNER JOIN groups g
ON u.id = g.uid
GROUP BY 1, 2;


-- Step 2: Creating convert_dt_agg
CREATE TEMPORARY TABLE convert_dt_agg AS
SELECT a.dt, g.group, AVG(a.spent) AS spending, COUNT(DISTINCT a.uid) AS
converted_user_count
FROM groups g
LEFT JOIN activity a
ON g.uid = a.uid
GROUP BY 1, 2;
```

```
--step 3
SELECT j.join_dt,j.group,
CAST(c.converted_user_count AS DECIMAL(10,2)) / j.user_count AS
convert_ratio,
c.sum
FROM join_dt_agg j
INNER JOIN convert_dt_agg c
ON j.join_dt = c.dt AND j.group=c.group;
```

3. **Python code for plotting the difference in spending and conversion,**

```
#import necessary packages
import matplotlib.pyplot as plt
import numpy as np

# Difference in means(data for two groups (A and B))
samplestatisticconv= 0.00706982258
samplestatisticspend= 0.0164
```

```
# Standard errors for the conversion rates and spendings
sd_diffconv= 0.00183
sd_diffspend= 0.2321

# Confidence interval calculation
confidence_interval1 = 1.96 * sd_diffconv  # 95% confidence interval
confidence_interval2 = 1.96 * sd_diffspend # 95% confidence interval

# Creating a combined bar chart with error bars with different colors
labels = ['Conversion Rate', 'Average Amount Spent']
values = [samplestatisticconv*100, samplestatisticspend*100]
conf_intervals = [confidence_interval1, confidence_interval2]
colors= [“green”, “pink”]

plt.bar(labels, values, yerr=conf_intervals, alpha=0.7, color=colors)
plt.title('Difference Comparison with Confidence Intervals')
plt.ylabel('Rate Difference(%)')
plt.show()
```

Click Here to the test calculator.

APPENDIx

- **Excel data work out**
- **Tableau data visualisations**