Introduction:

GloBox, an online marketplace known for its boutique fashion items and high-end decor products, has recently experienced significant growth in its food and drink offerings. To increase revenue and bring awareness to this expanding product category, GloBox's Growth team has decided to conduct an A/B test. This experimentation technique involves comparing two versions of the website to determine which one performs better in achieving specific goals. In this case, the test involves showcasing key products from the food and drink category as a banner at the top of the website. By implementing this A/B test, GloBox aims to evaluate the effectiveness of the banner in driving customer engagement and ultimately increasing sales.

Methodology:

SQL Querying:

- To obtain the necessary data for the analysis, SQL queries were executed to retrieve and combine information from multiple tables.
- The 'id' column from the 'users' table, the 'group' column from the 'groups' table, and the 'spent' column from the 'activity' table were selected as key variables of interest.
- A query was used to join the tables based on the common 'id' column, merging the user information with their corresponding group and spending data.
- The 'spent' values were summed up using an aggregate function to calculate the total spending for each user.
- Additionally, a boolean condition was applied to determine whether or not each user had spent within each group, allowing for further analysis of conversion rates.

*queries will be provided in appendix

Data Preparation:

- In the data preparation phase, after joining the tables and combining the data, it was observed that some users had missing values for their spending information. To ensure a consistent dataset for analysis, the null values in the spending column were replaced with zero (0). This step allowed for the inclusion of all users in the analysis, even if they had not recorded any spending. By replacing the null values with zero, we ensured

that the dataset was complete and ready for further analysis without excluding any users due to missing spending information.

Statistical Analysis with Python:

- Python, along with several libraries including pandas, numpy, scipy, seaborn, and matplotlib, was utilized for the statistical analysis.
- Hypothesis testing was conducted to compare the means between the control group (Group A) and the test group (Group B) using a t-test. A confidence interval of 95% was used, providing a range of values within which the true population mean difference is likely to fall.
- Another hypothesis test was performed to compare the proportions of conversions between the two groups using a Z-test. This allowed for the assessment of the statistical significance of the difference in conversion rates.
- Bar charts were created using matplotlib to visually represent the data, providing a clear visualization of the key metrics and comparisons between the control and test groups.
- The statistical analysis enabled a deeper understanding of the differences in means, proportions, and conversion rates between the groups, supporting the evaluation of the effectiveness of the banner and its impact on user behavior.

Significance Level:

- The significance level, also known as alpha, was set to 0.05 for the hypothesis tests conducted in this analysis.
- A significance level of 0.05 indicates that we are willing to accept a 5% chance of committing a Type I error, which is the rejection of a true null hypothesis. In other words, it sets the threshold for determining statistical significance.
- By using a significance level of 0.05, we are confident that any observed differences between the control and test groups that meet this threshold are unlikely to have occurred by chance alone.

Statistical Metrics:

- In terms of total spending, Group A accumulated a total of $82,145.90, whereas Group B had a higher total spending of $83,415.33. This demonstrates that users in Group B collectively spent more than those in Group A.
- After conducting a statistical analysis, it was determined that there is not enough evidence to support a significant difference in the average amount spent per user between the control group (Group A) and the treatment group (Group B). The average amount spent per user in Group A was $3.37, while in Group B it was slightly higher at $3.39. However, this difference was not found to be statistically significant at the chosen significance level of 0.05
- In contrast, the analysis revealed strong evidence to support a significant difference in the user conversion rate between the control group (Group A) and the treatment group (Group B). The conversion rate for Group A was 3.92%, whereas Group B exhibited a higher conversion rate of 4.63%. This difference in conversion rates was found to be statistically significant at the chosen significance level of 0.05, indicating that the inclusion of the banner in Group B had a notable impact on user behavior and led to a higher likelihood of conversion.

Conclusion:

Interpreting the results of the statistical analysis in light of our research objective sheds light on the effectiveness of the A/B test conducted by GloBox. The primary goal of the test was to bring awareness to the food and drink category and increase revenue for the company. Let's examine the findings:

- Average Amount Spent per User: Our analysis revealed that there was no significant difference in the average amount spent per user between the control group (Group A) and the treatment group (Group B). Although Group B had a slightly higher average spending per user ($3.39) compared to Group A ($3.37), this difference was not statistically significant. Therefore, we cannot conclude with confidence that the banner in Group B had a significant impact on the average spending per user.
- User Conversion Rate: On the other hand, the analysis showed strong evidence of a significant difference in the user conversion rate between the control group (Group A) and the treatment group (Group B). Group B exhibited a higher conversion rate of 4.63% compared to Group A's conversion rate of 3.92%. This statistically significant difference suggests that the inclusion of the banner in Group B positively influenced user behavior, resulting in a higher likelihood of conversion.

Considering the research objective of increasing revenue and driving conversions in the food and drink category, the A/B test results indicate that the banner in Group B played a vital role. It effectively captured users' attention and influenced their purchasing decisions, leading to a higher conversion rate compared to the control group.

Based on these findings, we recommend implementing the banner in the food and drink category site-wide to leverage its positive impact on conversions and overall revenue generation. Additionally, further analysis and optimization can be explored to maximize the effectiveness of the banner in driving user engagement and increasing sales in other product categories.

Recommendations:

- It's important to note that these conclusions are based on the data and analysis conducted during this specific A/B test period. Ongoing monitoring and further experimentation can provide additional insights into the long-term impact and potential variations in user behavior.
- Overall, the results of the statistical analysis provide valuable insights into the performance of the A/B test and its alignment with the research objective of increasing revenue in the food and drink category.

Appendix:

```
WITH cte AS (
  SELECT uid,
  "group",
  Sum(spent) AS total_spent
FROM groups
LEFT JOIN activity using(uid)
GROUP BY
```

```
  uid,
  "group"),
  cte_2 AS (
    SELECT
      uid,
      "group",
      (COALESCE(total_spent, 0)) total_spent
    FROM cte)
SELECT *,
CASE WHEN total_spent>0 THEN 1 ELSE 0 END AS converted
FROM cte_2
```