GLOBOX- A/B TESTING REPORT

COMPANY BACKGROUND:

GloBox is an e-commerce company. It is an online marketplace that specializes in sourcing unique and high-quality products from around the world.

PROJECT BACKGROUND:

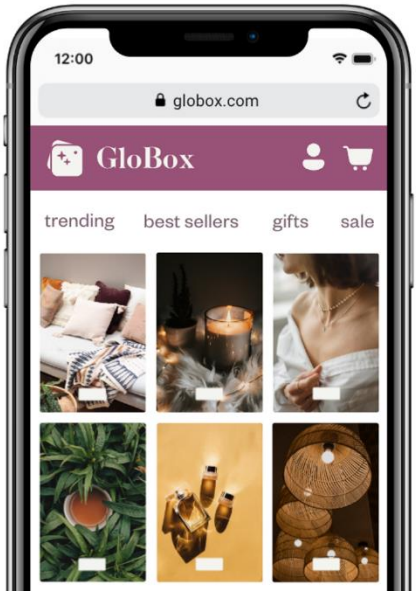
An A/B test is an experimentation technique used by businesses to compare two versions of a webpage, advertisement, or product feature to determine which one performs better. By randomly assigning customers or users to either the A or B version, the business can determine which version is more effective at achieving a particular goal.

GloBox is primarily known amongst its customer base for boutique fashion items and high-end decor products. However, their food and drink offerings have grown tremendously in the last few months, and the company wants to bring awareness to this product category to increase revenue.

The Growth team decides to run an A/B test that highlights key products in the food and drink category as a banner at the top of the website. The control group does not see the banner, and the test group sees it as shown below:

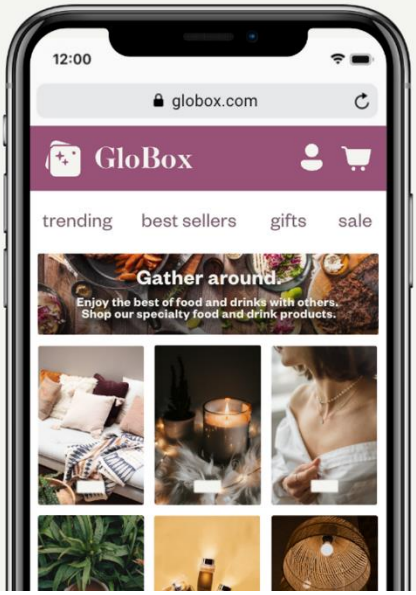
Group A: Control

existing landing page



Group B: Treatment

landing page with food & drink banner



PROJECT SETUP:

The setup of the A/B test is as follows:

1. The experiment is only being run on the mobile website.

2. A user visits the GloBox main page and is randomly assigned to either the control or test group. This is the join date for the user.
3. The page loads the banner if the user is assigned to the test group and does not load the banner if the user is assigned to the control group.
4. The user subsequently may or may not purchase products from the website. It could be on the same day they join the experiment, or days later. If they do make one or more purchases, this is considered a "conversion".

MY ROLE: I am working as Data Analyst for Glo-box e-commerce company. My task is to analyze the results of the A/B test and provide a recommendation to the stakeholders about whether GloBox should launch the experience to all users.

When comparing the performance of the control group (Group A) and the treatment group (Group B) in the A/B test, we can consider several metrics to measure the impact of the food and drink banner. Here are some metrics I compared:

1. CONVERSION RATE:

Compare the conversion rates of the control and treatment groups. Conversion rate refers to the percentage of website visitors who take a desired action, such as making a purchase. A higher conversion rate in the treatment group would indicate that the food and drink banner is effective in driving conversions.

	Control group (Group A)	Treatment group (Group B)
No. of conversions	955	1139
Total No. of users	24343	24600
Conversion Rate	**3.923099043**	**4.630081301**
Conversion Rate(in decimal)	0.03923099	0.046300813

NOTE: when spent> 0, it is considered as a conversion, please see attached excel sheet for the formula used.

By Using the below Formula:

Conversion Rate = (Number of Conversions / Total Number of Users) * 100

Control Group: 3.92%

Treatment Group: 4.63%

Comparing the user conversion rates between the control and treatment groups can provide insights into the effectiveness of the food and drink banner in driving conversions. In this case, the treatment group has a slightly higher conversion rate (4.63%) compared to the control group (3.92%). While the difference is not substantial, it suggests that the food and drink banner may have a positive impact on conversion rates.

Hypothesis Testing: *Why is it required?*

Conducting further analysis and hypothesis testing can help determine if the observed difference is statistically significant and if it can be attributed to the food and drink banner or other factors.

Now we are conducting a hypothesis test to see whether there is a difference in the conversion rate between the two groups.

Conduct the Hypothesis Test: Use the two-sample z-test for proportions, assuming equal proportions.

Calculate the Test Statistic and P-value: Calculate the test statistic and p-value using the appropriate statistical software or calculator. The test statistic is typically the z-value, and the p-value represents the probability of observing the data given the null hypothesis.

Evaluate the Results: Compare the p-value to the significance level. If the p-value is less than the significance level, reject the null hypothesis. Otherwise, if the p-value is greater than or equal to the significance level, fail to reject the null hypothesis.

Formulate the Hypotheses:

Null Hypothesis (H0): There is no difference in the user conversion rate between the control and treatment groups.
Alternative Hypothesis (H1): There is a difference in the user conversion rate between the control and treatment groups.
Set the Significance Level: In this case, the significance level is 0.05, corresponding to a 5% significance level.

We are using a two-sample two-sided z-interval for a difference in proportions. Assuming equal proportions, we use the pooled standard error.

What is the resulting p-value and conclusion? Use the normal distribution and a 5% significance level. Use the pooled proportion for the standard error.

See excel sheet (question 9) for Calculations:

pooled proportion	0.042784		(converted_usersA+converted usersB)/(sample_sizeA+sample_sizeB)	Pooled proportion (p) = (Number of conversions (control) + Number of conversions (treatment)) / (Total number of users (control) + Total number of users (treatment))
standard error	0.001830		Standard Error = sqrt((pooled proportion * (1 - pooled proportion)*(1/sample_sizeA+1/sample_sizeB)	Standard error (SE) = sqrt(p * (1 - p) * (1 / Total number of users (control) + 1 /

				Total number of users (treatment)))
Z-score	3.864292		(convertion rate B-convertion rate A)/standard error	
p value	**0.000111**		**p-value = 2 * (1 - NORM.S.DIST(3.864, TRUE))**	

My answer: p = 0.0001, statistically _significant_. We _reject_ the null hypothesis that there is no difference in the user conversion rate between the control and treatment.

With a p-value of 0.0001, which is less than the significance level of 0.05, the resulting conclusion is that the difference in the user conversion rate between the control and treatment groups is statistically significant. Therefore, we reject the null hypothesis. This suggests that there is a significant difference in the user conversion rate between the two groups based on the data obtained during the A/B test.

Insight Gained:

The statistical significance in the difference of the user conversion rates implies that the food and drink banner likely had an impact on driving conversions. The treatment group, which saw the banner, exhibited a significantly higher conversion rate compared to the control group. This suggests that the food and drink banner was effective in influencing users to convert and engage with the offerings in that category.

This insight can inform decision-making and further optimization efforts. The results indicate that highlighting key products in the food and drink category through the banner has a positive impact on the user conversion rate. Therefore, it may be beneficial to continue promoting and refining the food and drink offerings to drive even higher conversion rates and revenue growth.

2. AVERAGE AMOUNT SPENT PER USER

Conduct a hypothesis test to see whether there is a difference in the average amount spent per user between the two groups. What are the resulting p-value and conclusion?

Use the t distribution and a 5% significance level. Assume unequal variance.

We are using a two-sided t-test for a difference in means. Assuming unequal variance, we use the unpooled standard error.

To conduct a hypothesis test to determine if there is a difference in the average amount spent per user between the control group (Group A) and the treatment group (Group B), we can follow these steps:

Formulate the Hypotheses:

Null Hypothesis (H0): There is no difference in the mean amount spent per user between the control and treatment groups.
Alternative Hypothesis (H1): There is a difference in the mean amount spent per user between the control and treatment groups.
Set the Significance Level: In this case, the significance level is 0.05, corresponding to a 5% significance level.

Conduct the Hypothesis Test: Use the two-sample t-test for independent samples, assuming unequal variance.

Calculate the Test Statistic and P-value: Calculate the test statistic and p-value using the appropriate statistical software or calculator. The test statistic is typically the t-value, and the p-value represents the probability of observing the data given the null hypothesis.

Evaluate the Results: Compare the p-value to the significance level. If the p-value is less than the significance level, reject the null hypothesis. Otherwise, if the p-value is greater than or equal to the significance level, fail to reject the null hypothesis.

See excel sheet (question 4) for Calculations:

p-value		
Significance level: 5%	0.05	
p-value	**0.943855753**	**=T.TEST (range1, range2, tails, type)**

My answer: p = 0.944, statistically insignificant. We fail to reject the null hypothesis that there is no difference in the mean amount spent per user between the control and treatment.

With a p-value of 0.944, which is greater than the significance level of 0.05, the resulting conclusion is that the difference in the average amount spent per user between the control and treatment groups is statistically insignificant. Therefore, we fail to reject the null hypothesis. This suggests that there is no significant difference in the mean amount spent per user between the two groups based on the data obtained during the A/B test.

Insight Gained:
The insight we gain from the above analysis is that there is no significant evidence to suggest that the treatment group had a different mean amount spent per user compared to the control group. The p-value of 0.944, which is greater than the significance level of 0.05, indicates that the observed difference in means is likely due to chance variation rather than a true difference between the groups. **This implies that the treatment, whatever it may be, did not have a noticeable effect on the amount spent per user. It suggests that the control and treatment groups behaved similarly in terms of their spending habits, and any variations observed could be attributed to random fluctuations or other factors not related to the treatment. Therefore, based on this analysis, there**

is insufficient evidence to conclude that the treatment had a significant impact on the average amount spent per user.

3. CONFIDENCE INTERVALS

CI for each group : See excel sheet (question 2 & 3) for Calculations:

	Control group (Group A)	**Treatment group (Group B)**	
Standard deviation	25.93639056	25.4141096	
Count/sample size	24343	24600	
Standard error	0.166235008	0.162034445	**= stdev/SQRT(sample size)**
degree of freedom	24342	24599	**=sample size-1**
T Critical value	1.960061445	1.960060427	**=T.INV.2T(0.05, degrees of freedom)**
The Margin of Error	0.325830829	0.317597303	**=T critical value * Standard Error**
Confidence Intervals(lower)	**3.048687639**	**3.073269643**	**=mean - margin of error**
Confidence Intervals(lower)	**3.700349297**	**3.708464249**	**=mean+ margin of error**

The resulting confidence interval for the average amount spent per user in the control group is (3.049, 3.700). The resulting confidence interval for the average amount spent per user in the treatment group is (3.073, 3.708).

Insight Gained:

Overlapping intervals: The confidence intervals for the average amount spent per user in the control group (3.049, 3.700) and the treatment group (3.073, 3.708) overlap. This indicates that there is no statistically significant difference in the average amount spent per user between the control and treatment groups.

Similar range: The range of the confidence intervals is quite close in both groups. The lower bounds (3.049 and 3.073) and the upper bounds (3.700 and 3.708) are relatively similar. This suggests that the average amount spent per user in both groups is within a similar range.

Based on these insights, it appears that the introduction of the food and drink banner in the treatment group did not have a significant impact on the average amount spent per user compared to the control group. The confidence intervals overlap, indicating that the difference in spending behavior between the two groups is not statistically significant.

OTHER METRICS

It's important to consider other factors and metrics in conjunction with the above metrics to gain a comprehensive understanding of the impact of the food and drink banner. Analyzing click-through rates, or other user engagement metrics can provide additional insights into the effectiveness of the banner in driving user behavior and revenue. For example,

1. Click-through Rate (CTR): If the food and drink banner is a clickable element, you can measure the CTR in the treatment group. A higher CTR would indicate that the banner is capturing users' attention and generating interest in the food and drink category.
2. Engagement Metrics: Consider metrics related to customer engagement with the food and drink category, such as time spent on the category page, the number of products viewed, or the number of add-to-cart actions. Higher engagement in the treatment group would suggest that the banner is effectively drawing attention to the food and drink offerings.
3. Revenue: This is the primary metric to evaluate the effectiveness of the food and drink banner. Compare the revenue generated by the control group (without the banner) with the treatment group (with the banner) to see if there is a significant difference.

These metrics can provide valuable insights into the impact of the food and drink banner on revenue and customer behavior. It's important to choose metrics that align with the specific goals of GloBox and reflect the desired outcomes of the A/B test.

Additionally, it's essential to monitor these metrics over time and consider the potential long-term effects of the food and drink offerings on user behavior and revenue. A single A/B test may not capture the full impact, and continuous optimization and experimentation may be required to drive significant changes in user spending patterns.