Addressing Membership Challenges: Insights and Strategies for Amour Beauty Box

Khanh Nguyen, Victoria Simon, Mykhailo Kuian

Objective

Amour Beauty Box, a subscription service specializing in indie beauty brands, has seen declining new sign-ups despite strong retention. Surveys and A/B testing revealed unclear pricing, lack of reviews, and missing product previews as barriers, while first-time discounts and clearer product details boosted interest. To attract more subscribers, we recommend optimizing the "Join Now" button, integrating product previews earlier, and testing first-time discounts.

Sample size and duration of the experiment

Last year, Amour Beauty Box reportedly had around 30,000 subscribers and received approximately 15,000 website visits per month, which averages to about 500 visits per day.

The estimated proportion is calculated as the number of subscribers divided by the total visits.

$$p = \frac{30{,}000 \text{ subscribers}}{15{,}000 \text{ visits per month } \times 12 \text{ months per year}} \approx 0.167$$

We want to construct a 95% confidence interval for p = 16.7% with a margin of error M = 1.3%. The minimum sample size is:

$$n = \left(\frac{z}{M}\right)^2 \cdot p(1-p) = \left(\frac{1.960}{0.013}\right)^2 \cdot 0.167 \cdot (1 - 0.167) \approx 3{,}162.2$$

We aimed to collect at least 3,162 samples for each group. Additionally, to account for duplicate users accessing the website, we set a goal of collecting 10,000 samples.

The duration of the experiment is calculated by dividing the number of samples that need to be collected by the number of daily visits.

$$\text{Duration} = \frac{10{,}000 \text{ samples}}{500 \text{ visits per day}} = 20 \text{ days}$$

We will test changes to the website and signup flow to evaluate their impact on conversion rates. The control and variant groups will each consist of 3,162 samples (6,324 total, accounting for duplicates, aiming for 10,000 samples overall). The test will run for approximately 20 days.

Testing metric

There are two ways to conduct an A/B test, depending on the metric used:

- **Conversion** refers to any action taken online that aligns with a business's goals, such as filling out a form, making a purchase, or completing a survey.
- **Revenue** is a continuous metric that directly impacts the business's bottom line.

Conversion rate measures the percentage of visitors who complete a desired action, helping businesses evaluate website effectiveness. A/B testing on revenue clarifies that a high conversion rate doesn't always mean higher earnings—Amour Beauty Box, for instance, may see more sign-ups for its monthly plan, increasing conversions but lowering revenue, whereas fewer annual subscriptions could yield higher revenue despite a lower conversion rate.

Design A/B testing

To assess the impact of signup flow improvements at Amour Beauty Box, we tested two hypotheses—one on conversion rate and another on revenue per user—based on the placement of the "Join Now" button. The control design positioned the button in the top left corner, while the variant placed it centrally for better visibility and engagement.

For **conversion rate:**

- N**ull hypothesis (H_{01})** asserted that the proportion of users signing up was the same between the control group and the variant group. That is, the change implemented in the variant group does not lead to a statistically significant increase in conversion rate.
- **Alternative hypothesis (H_{11})** proposed that signup rates differed between the two groups, indicating that the implemented change has led to a statistically significant increase in conversion rate.

Similarly, for **revenue**:

- **Null hypothesis (H_{02})** posited that there was no difference in revenue per user between the control and variant groups
- **Alternative hypothesis (H_{12})** suggested that revenue differs across groups.

Figure. (Left)The control group, where the 'Join Now' button is positioned at the top left. (Right) The variant group, where the ‘Join Now’ button is positioned at the center.

Data Analysis

Over 20 days, we collected 10,000 samples and removed duplicate users to prevent skewed results. Due to non-normal data distributions, we used Mann-Whitney U-tests instead of Z-tests to compare signup behavior and revenue. The p-values for both metrics exceeded 0.05, leading us to fail to reject the null hypotheses. Full analysis details are available on GitHub. Full analysis details are available on GitHub.

Conclusion

The implemented change did not significantly impact sign-ups or revenue, as we failed to reject both null hypotheses. This suggests that conversion rates and purchasing behavior remain influenced by other factors. Further analysis is needed to explore user preferences, market conditions, pricing strategies, and website optimizations that could enhance engagement and financial growth.