

The A/B Testing of International Online Shop

By: Diana Septiani



General Conclusions

There are differences between the A/B testing results for statistical differences between the proportions using Z-test

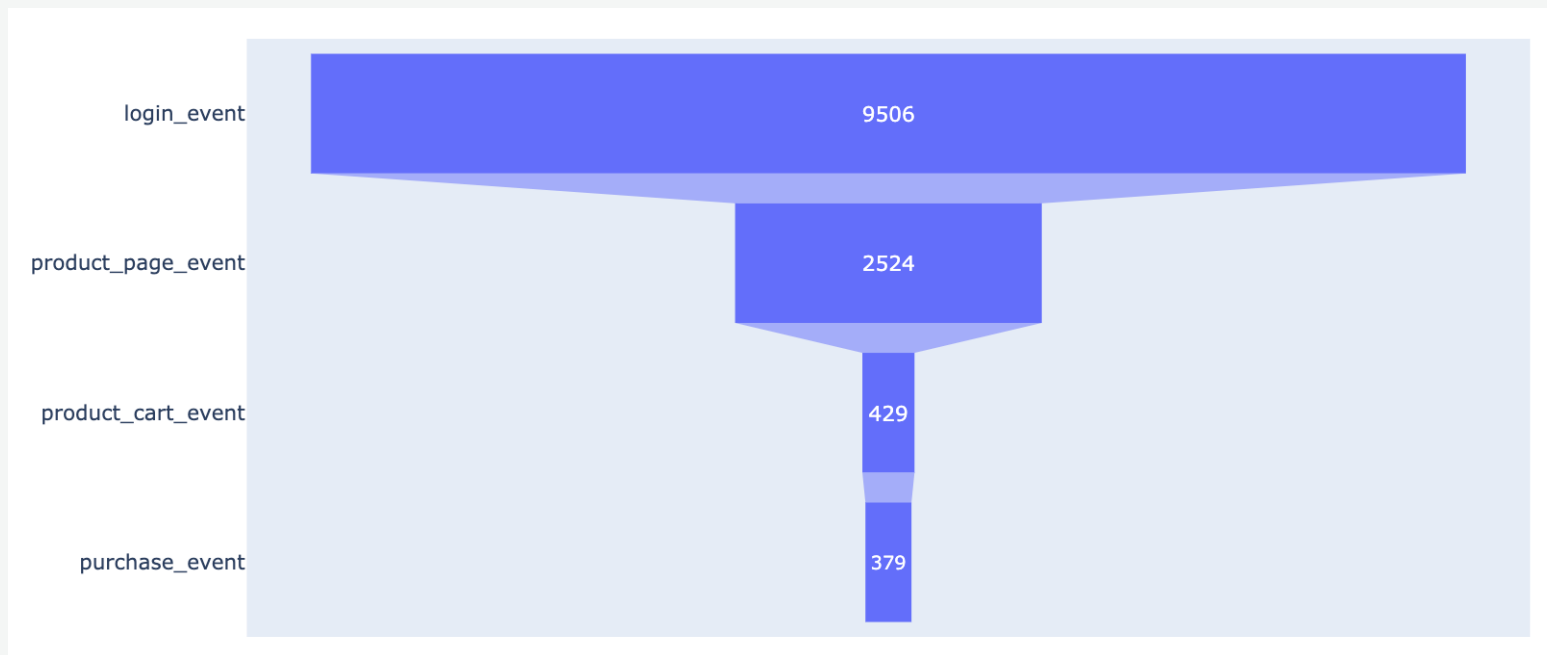
- "alpha" value = 0.01 → failed to reject the null hypothesis
- "alpha" value = 0.05 → reject the null hypothesis

There is a possibility of a significant difference between the shares of group samples A and B. Because decreasing alpha from 0.05 to 0.01 makes it harder to reject the null hypothesis.

We assume that the hypothesis test in group B has no effect on the increase in events because there are only a few users in group B.

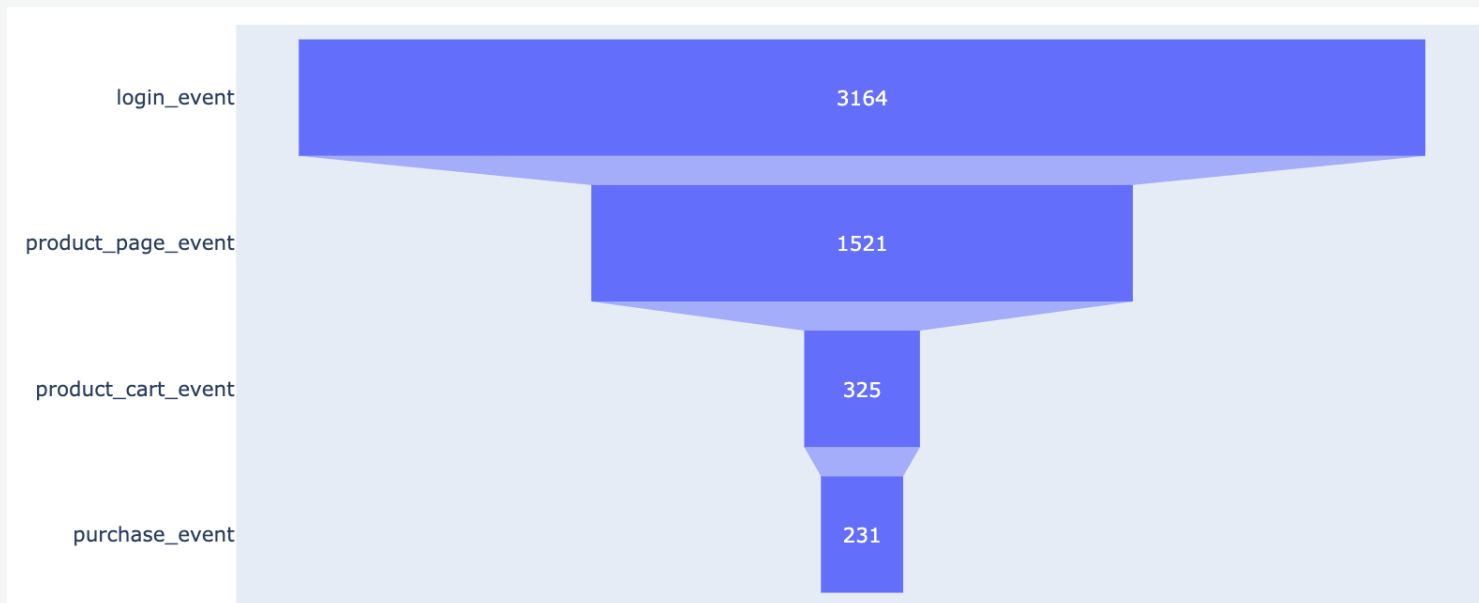
Funnel by Total Events

By the Closed Funnel approach, we can see that from 9506 login events, only 2524 continue to the product_page, 429 events continue to the 'product_cart, and 379 events are completed until the purchase stage.



Funnel by Total Users

- We have 3164 users who have performed on the event of login, with the detail as follows:
- 1521 users have performed on the event of product_page.
- 325 users have performed on the event of purchase.
- 231 users have performed on the event of product_cart.
- The percentage of users who make the entire journey from the first event to the checkout stage is 7.3% (3163 users vs. 231 users).



The average distribution of events per user is 77.3% in Group A and 22.6% in Group B. It's evenly distributed for each event.

group	A	B	total	%A	%B
log_ev_user	2330	834	3164	0.736410	0.263590
prod_page_ev_user	1176	345	1521	0.773176	0.226824
prod_cart_ev_user	258	67	325	0.793846	0.206154
purc_ev_user	183	48	231	0.792208	0.207792

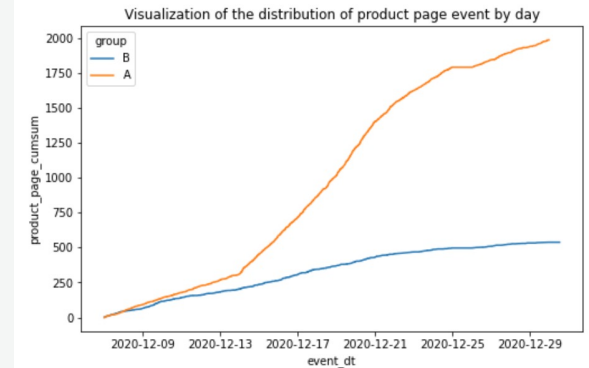
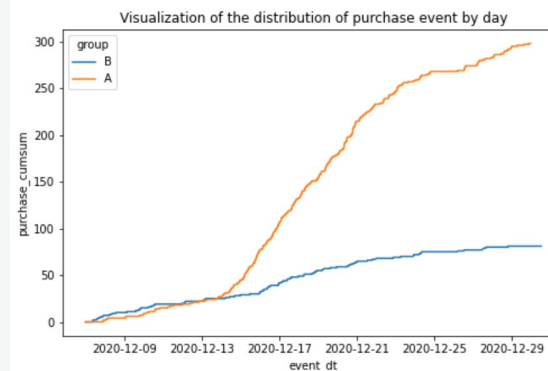
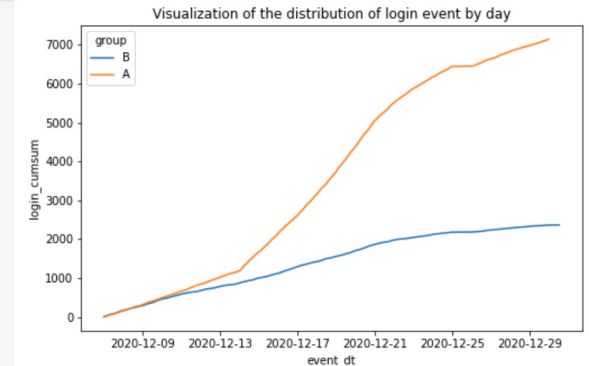
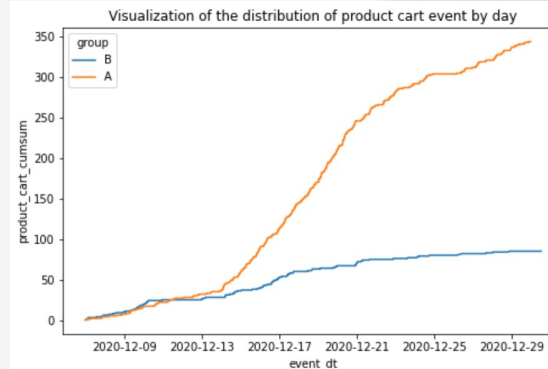
There are no users double in both samples.


- From the description, the total group_a is 2330 users, and group_b is 834 users, equal to 3164 users in column user_id.

total_users	
A	2330
B	834

Group A generates more events than group B.

- Group A generates more events than Group B.
- This may be because the number of users in Group A is far higher than in Group B (2330 users vs. 834 users).
- It can be assumed that the hypothesis test in group B has no effect on the increase in events because there are only a few users.





The hypothesis testing below, with the significance level / "alpha" value to be used = 0.01, the result is failed to reject the null hypothesis.


- There is **no statistically significant** shares difference between group samples A and B, except for the event of product_page.


```
total_users    1.0
Name: A, dtype: float64 event log_ev_user group A
total_users    1.0
Name: B, dtype: float64 event log_ev_user group B
p-value: [nan]
Failed to reject the null hypothesis, there is no reason to consider the shares different

total_users    0.504721
Name: A, dtype: float64 event prod_page_ev_user group A
total_users    0.413669
Name: B, dtype: float64 event prod_page_ev_user group B
p-value: [6.2922157e-06]
Reject the null hypothesis: there is a significant difference between the shares

total_users    0.11073
Name: A, dtype: float64 event prod_cart_ev_user group A
total_users    0.080336
Name: B, dtype: float64 event prod_cart_ev_user group B
p-value: [0.01309844]
Failed to reject the null hypothesis, there is no reason to consider the shares different

total_users    0.078541
Name: A, dtype: float64 event purc_ev_user group A
total_users    0.057554
Name: B, dtype: float64 event purc_ev_user group B
p-value: [0.04558298]
Failed to reject the null hypothesis, there is no reason to consider the shares different
```





The hypothesis testing below, with the significance level / "alpha" value to be used = 0.05, the result is reject the null hypothesis.

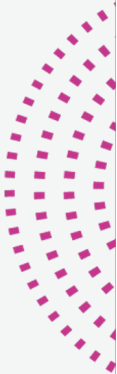
- There is a **significant difference** between the shares of group samples A and B, except for the event of login.

```
total_users    1.0
Name: A, dtype: float64 event log_ev_user group A
total_users    1.0
Name: B, dtype: float64 event log_ev_user group B
p-value: [nan]
Failed to reject the null hypothesis, there is no reason to consider the shares different

total_users    0.504721
Name: A, dtype: float64 event prod_page_ev_user group A
total_users    0.413669
Name: B, dtype: float64 event prod_page_ev_user group B
p-value: [6.2922157e-06]
Reject the null hypothesis: there is a significant difference between the shares

total_users    0.11073
Name: A, dtype: float64 event prod_cart_ev_user group A
total_users    0.080336
Name: B, dtype: float64 event prod_cart_ev_user group B
p-value: [0.01309844]
Reject the null hypothesis: there is a significant difference between the shares

total_users    0.078541
Name: A, dtype: float64 event purc_ev_user group A
total_users    0.057554
Name: B, dtype: float64 event purc_ev_user group B
p-value: [0.04558298]
Reject the null hypothesis: there is a significant difference between the shares
```



Recommendations



It's better to increase the no of observation population and longer the experiment days to gain more convincing A/B testing results.

Because, the hypothesis test in group B has no effect on the increase in events because there are only a few users in group B.

There is a possibility of a significant difference between the shares of group samples A and B. Because decreasing alpha from 0.05 to 0.01 makes it harder to reject the null hypothesis.

Thank you

