

Analysis of mobile application A/B test results

Done: **Vladislav Goncharuk**

Goal: Determine whether a "50% off" indication affects subscription conversions.

Tools: Python (pandas, scipy, seaborn, matplotlib)

Project done in course Data Analytics from **GoIT School**

pandas

seaborn

Data description

Data:

One row= one user

Fields: user_id, test_group (a/b), conversion (0/1), event_time

Goal: Compare conversion (install → purchase) between groups

Group	Users	Conversion
A	10 013	6.10%
B	9 985	8.90%

Data quality check:

- ✓ No empty values
- Unique users: **19 998**
- Test start: **03.07.2023**
- Test end: **25.07.2023**
- Duration: **23 days**

```
test_group
a      10013
b       9985
Name: user_id, dtype: int64
test_group
a      0.061021
b      0.089034
Name: conversion, dtype: float64
```

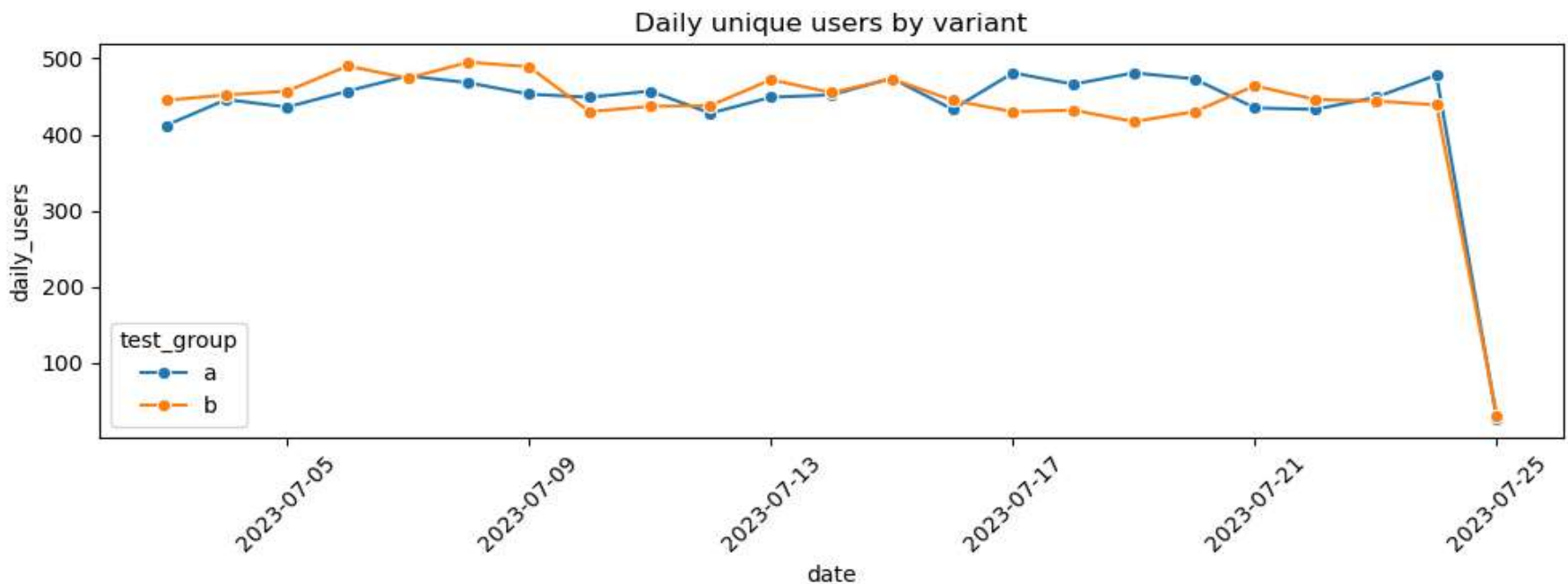
pandas

seaborn

User dynamics

Conclusion: Traffic is stable, no anomalies observed.

On chart: Axis X — date, Axis Y — daily users count



The line graph shows the number of unique users by day for groups A and B.

Statistical analysis

Method: Since the metric is binary (conversion), a **z-test for two options** was applied, **χ^2 -test** and **permutation test**.

α - 5%

	Test	Statistic	p-value	Conclusion
	Z-test	-7.52	0.0000000000000549	Reject H_0
Conversion in group B is statistically significantly higher than in group A.	χ^2 -test	56.14	0.0000000000000674	Reject H_0
	Permutation	-0.028	0.0002	Reject H_0

Test results:

```
=== Summary ===
      conversions      n      cr
test_group
a              611  10013 0.0610
b              889   9985 0.0890

=== Statistical Tests ===
Z-test: z= -7.520, p=0.00000000000005491218 → Reject  $H_0$  - option B wins.
Chi2-test:  $\chi^2$ = 56.142, p=0.00000000000006740441
Permutation test: stat= -0.028, p=0.000200
```

Пояснення: p-value < 0.05 → відхиляємо H_0

Conversion rate compare

Results:

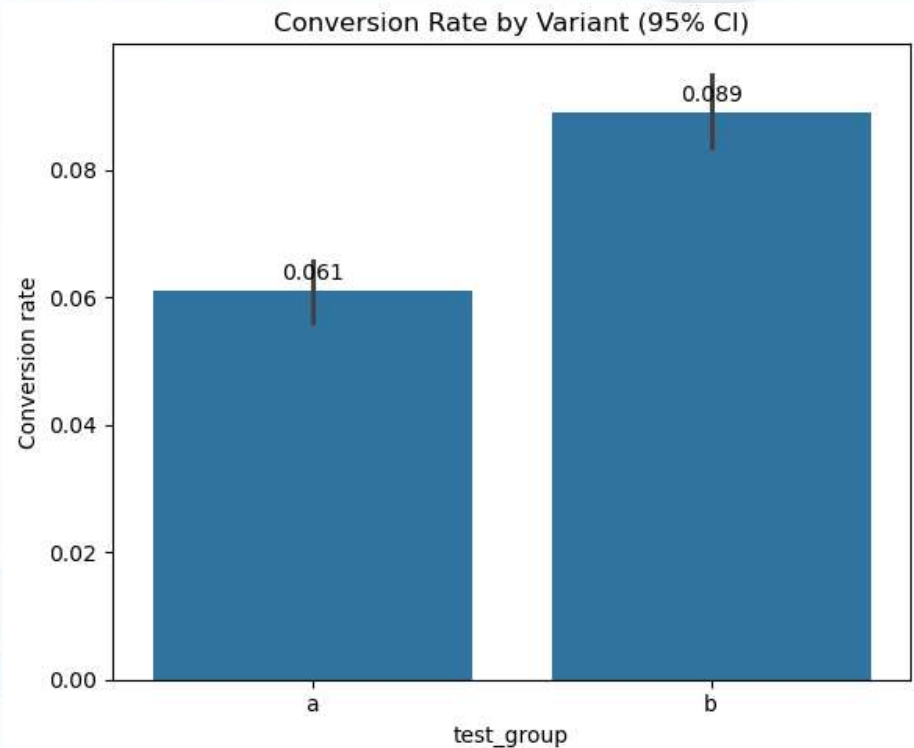
absolute difference

$\Delta_{\text{abs}} (B-A) = +2.8$ percentage points.

Relative growth (Uplift) = **+45.9%**

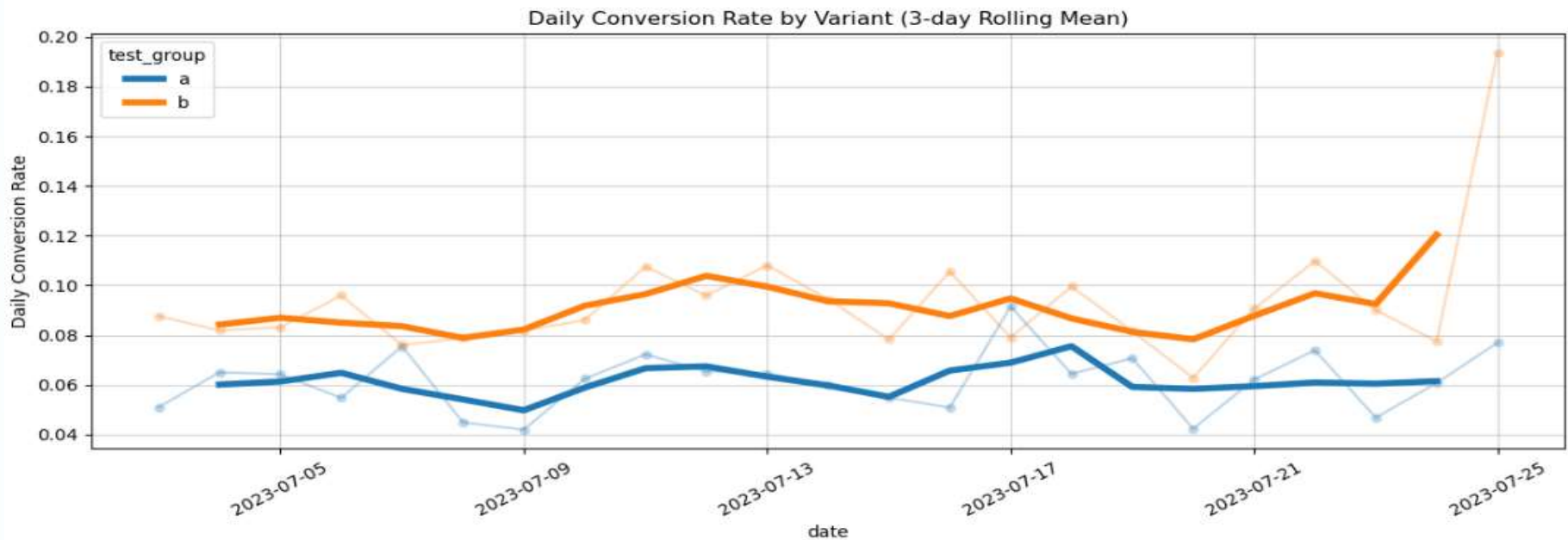
	conversions	n	cr
test_group			
a	611	10013	0.061
b	889	9985	0.089

$\Delta_{\text{abs}}(B-A)=0.0280$, uplift=45.91%



Graph with 95% confidence intervals - **CR difference between A and B.**

Conversion dynamics



Results: The effect is not random, the trend is consistent.

Conclusion: The chart of daily CR for each group demonstrates a stable, non-random advantage of Variant B throughout the entire testing period. Trend lines are also shown to smooth data variability.

pandas

seaborn

Recomendations

Results: The effect is not random, the trend is consistent.

Value	A	B
User count	10 013	9 985
Conversion Rate	6.10%	8.90%
Δ abs	+2.8 p.p.	
Uplift	+45.9%	

```
N_A: 10013
CR_A: 0.061
N_B: 9985
CR_B: 0.089
 $\Delta$ abs(B-A): 0.0280
Uplift: 45.91%
p_z: 0.000000000000005491218
p_chi2: 0.000000000000006740441
p_perm: 0.000200
```

Conclusion: The difference is **statistically significant** (p-value < 5%).

- ✓ **Recommendation:** Implement a new design with a "50% discount" tag.
- ✓ **Additionally** — check the impact on **LTV** and **user churn**.

Thanks for attention!

Ready to cooperate

Contacts

 [Email](#)

 [LinkedIn](#)

 [GitHub](#)

pandas

seaborn