

Завдання:

Отримали результати а/б тесту перших сторінок сайту, за яким кожен п'ятий користувач потрапляв на нову тестову сторінку.

- А (базовий варіант взаємодії): **реklamний креатив → сайт з дефолтною сторінкою**
- В (тестовий варіант): **реklamний креатив → сайт з новою першою сторінкою**

Усі наступні сторінки сайту, офери, фінальна ціна та рекламні креативи залишалися незмінними.

Запитання:

- Чи можна оцінити ефективність лендингів, базуючись на цих даних?
- Які наступні дії потрібно зробити команді?

Спробую оцінити ефективність лендингів, орієнтуючись на отримані дані. Для цього додамо нові метрики у нашу таблицю:

CPC (Cost Per Click) - ціна за клік

CPCon (Cost Per Conversion) - ціна за цільову дію користувача

Conversion(%) – відсоток користувачів, які виконали бажану дію

Також додам відсоток зміни метрики у тестовому варіанті порівняно з базовим (change percentage). Цей показник будемо додавати лише до нових метрик.

	Click	Conversion	Budget	CPC	CPCon	Conversion(%)
Base	1630188	3163	133802.85	0,08208	42,3	0,00194
Test	418453	857	40906.25	0,09776	47,73	0,002048
Change percentage	-	-	-	+19,1%	+12,8%	+5,6%

Як можна побачити, відсоток конверсії збільшився (+5,6%) на тестовому варіанті, що є добре для продукту, проте значно виросли метрики вартості кліку та цільової дії користувача (+19,1% та +12,8%). Відповідно, у разі переходу на тестовий варіант, виросте бюджет, необхідний для реклами. Порівнюючи збільшення конверсії та бюджету, можна прийти висновку, що краще залишитись на базовому варіанті.

Проте, нам важливо перевірити статистичну значимість отриманого результату. Тобто впевнитись, чи не випадкові значення ми отримали і чи достатній об'єм вибірки для того, щоб робити висновок базуючись на ці дані.

Щоб перевірити значимість отриманого результату, скористаємось [онлайн калькулятором для а/б тесту](#):

Calculate your statistical significance

	Visitors	Conversions	Conversion rate
A	1630188	3163	0.19%
B	418453	857	0.20%

Hypothesis ☒ One-sided ☐ Two-sided Confidence ☐ 90% ☒ 95% ☐ 99%

Calculate

Result not significant!

Variant B's conversion rate (0.20%) was 5.55% higher than variant A's conversion rate (0.19%), but you cannot say, with 95% confidence, that variant B will perform better than variant A.

Power 76.75% **p value** 0.0834

Бачимо, що даних не достатньо для того, щоб зробити висновок про ефективність лендингів. Знайдемо мінімальний розмір вибірки, щоб уникнути не значущих результатів, за допомогою [онлайн калькулятору](#) (помилка 1 роду 0,05, помилка другого роду 0,2):

Question: How many subjects are needed for an A/B test?

Baseline conversion rate: % [link]

Minimum Detectable Effect: % 0.19% – 0.21%

The Minimum Detectable Effect is the smallest effect that will be detected (1-β)% of the time.

☒ Absolute ☐ Relative

Conversion rates in the gray area will not be distinguishable from the baseline.

Sample size:
2,707,918
per variation

Statistical power 1-β: % Percent of the time the minimum effect size will be detected, assuming it exists

Significance level α: % Percent of the time a difference will be detected, assuming one does NOT exist

Отже, для отримання значущого результату треба мінімум збільшити об'єм вибірки до 2707918.

Висновок:

Чи можна оцінити ефективність лендингів, базуючись на цих даних?

Ні. Результат не є статистично значимим при значення помилки першого роду 0,05 і помилки другого роду 0,2.

Які наступні дії потрібно зробити команді?

Команді треба зібрати більше даних на тестовому варіанті, щоб уникнути статистично не значимого результату.