

Marketing A/B Testing: Оценка эффективности рекламной кампании

Kaggle FavioVázquez Dataset

Трусов Андрей

Kaggle Data Analysis

2025

Содержание

1	Введение	3
2	Методология	3
2.1	Дизайн эксперимента	3
2.2	Проверки корректности	4
2.3	Методы статистического анализа	4
2.3.1	Z-тест для пропорций	4
2.3.2	Бутстреп для доверительного интервала Lift	5
2.3.3	Анализ почасовых паттернов	5
2.3.4	Анализ зависимости от частоты показов	5
3	Результаты	6
3.1	Основные метрики	6
3.2	Основная гипотеза	6
3.3	Lift (Относительный прирост)	6
3.4	Анализ почасовых паттернов	8
3.4.1	Ключевые выводы по часам	8
3.5	Анализ зависимости от частоты показов (Frequency Capping)	9
3.5.1	Тренд конверсии по бакетам	9
3.5.2	Интерпретация	9
4	Результаты	10
4.1	Глобальные показатели эффективности	10
4.2	Статистическая значимость	10
4.3	Оценка Lift и доверительные интервалы	11
4.4	Почасовая эффективность	11
4.5	Анализ частоты (Frequency Capping)	12

1 Введение

Данный отчет представляет результаты анализа A/B теста маркетинговой кампании, проведенного на датасете Marketing A/B Testing (Kaggle, FavioVázquez). Целью исследования является:

1. Оценить статистическую и практическую значимость влияния рекламы на конверсию
2. Выявить временные (часовые) и частотные паттерны эффективности
3. Предоставить рекомендации по оптимизации рекламной кампании

Объем выборки составляет 588 101 уникальных пользователей, что обеспечивает высокую статистическую мощность теста.

2 Методология

2.1 Дизайн эксперимента

Эксперимент построен по классической схеме A/B теста:

- Группа А (Treatment, “ad”): 564 577 пользователей (96%), которым показывалась целевая реклама
- Группа В (Control, “psa”): 23 524 пользователя (4%), которым показывалась социальная реклама (PSA) вместо целевой рекламы

2.2 Проверки корректности

Перед началом анализа выполнены следующие проверки:

1. Проверка дубликатов: пользователи, попавшие одновременно в обе группы (cross-contamination).
2. Проверка пропусков: Отсутствуют пропуски в критических полях (user_id, test_group, converted).

2.3 Методы статистического анализа

2.3.1 Z-тест для пропорций

Для тестирования основной гипотезы применен Z-тест для двух независимых пропорций:

$$Z = \frac{p_1 - p_2}{\sqrt{p(1-p) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (1)$$

где:

- p_1, p_2 — доли конверсий в группах А и В
- $p = \frac{x_1+x_2}{n_1+n_2}$ — объединенная доля
- n_1, n_2 — размеры групп

Гипотезы:

- $H_0 : p_{ad} = p_{psa}$ (нет различий)
- $H_1 : p_{ad} \neq p_{psa}$ (различия существуют, двусторонний тест)

Уровень значимости: $\alpha = 0.05$.

2.3.2 Бутстреп для доверительного интервала Lift

Для оценки устойчивости показателя Lift применена техника bootstrap:

$$\text{Lift} = \frac{CR_{ad} - CR_{psa}}{CR_{psa}} \times 100\% \quad (2)$$

где CR — Conversion Rate (доля конвертировавшихся).

Процедура:

1. Из каждой группы проведено 2000 ресэмплирований с возвращением
2. На каждом ресэмплировании рассчитана новая величина Lift
3. Построен эмпирический доверительный интервал на уровне 95% (2.5% и 97.5% квантили)

2.3.3 Анализ почасовых паттернов

Для выявления временных закономерностей вычислена конверсия по часам суток:

$$CR_{h,g} = \frac{\sum_{i=1}^n \text{converted}_i \cdot 1_{h_i=h} \cdot 1_{g_i=g}}{\sum_{i=1}^n 1_{h_i=h} \cdot 1_{g_i=g}} \quad (3)$$

где h — час суток (0–23), g — группа (ad или psa).

2.3.4 Анализ зависимости от частоты показов

Пользователи разбиты на бакеты в зависимости от количества увиденной ими рекламы:

- 1–10 показов
- 11–20 показов
- 21–50 показов
- 51–100 показов
- 101–150 показов
- 151–200 показов

- 201–300 показов
- 301–500 показов
- 500+ показов

Для каждого бакета рассчитана конверсия в группе “ad”, что позволяет выявить эффект насыщения.

3 Результаты

3.1 Основные метрики

Метрика	Группа AD	Группа PSA
Количество пользователей	564 577	23 524
Количество конверсий	14 423	420
Conversion Rate	2.55%	1.79%

Таблица 1: Основные метрики по группам

3.2 Основная гипотеза

Гипотеза: Показ целевой рекламы приводит к статистически значимому увеличению конверсии по сравнению с контрольной группой (PSA).

Z-тест результаты:

- Z-statistic ≈ 7.37
- P-value < 0.0001

Вывод: H_0 отвергаем. Различие в конверсии между группами статистически значимо на уровне 0.05.

3.3 Lift (Относительный прирост)

Глобальный Lift:

$$\text{Lift} = \frac{0.025547 - 0.017854}{0.017854} \approx 0.4309 \approx 43.09\%.$$

(4)

Bootstrap доверительный интервал (95%):

- Нижняя граница: $\approx 30.25\%$
- Верхняя граница: $\approx 58.13\%$
- Средний Lift по бутстрепу: $\approx 43.55\%$

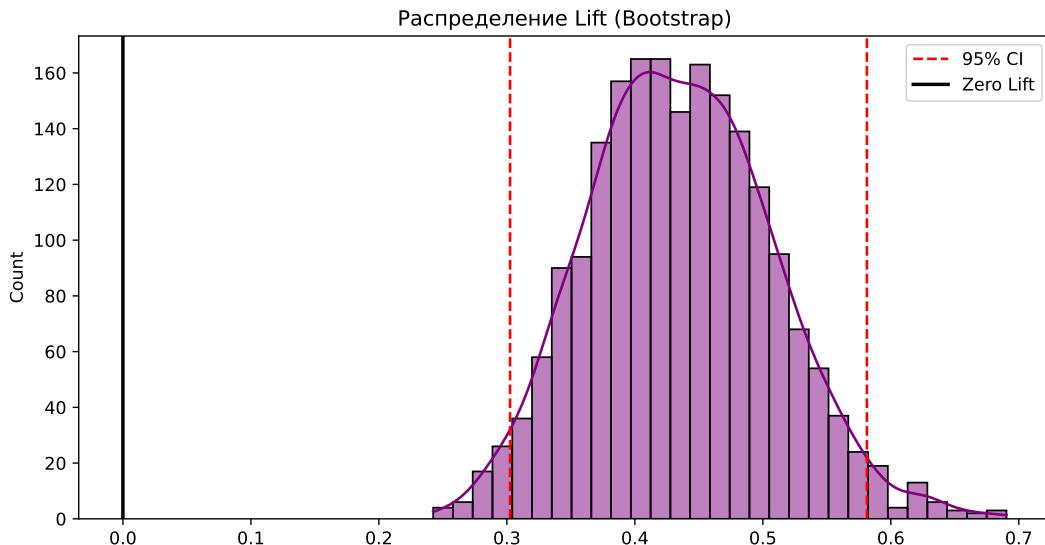


Рис. 1: Гистограмма распределения Lift, посчитанного методом Bootstrap. Красными линиями отмечены 2.5 и 97.5 процентили, черная сплошная линия - уровень нулевого эффекта.

Интерпретация: С вероятностью 95% мы уверены, что истинный Lift лежит в диапазоне от 30.25% до 58.13%.

3.4 Анализ почасовых паттернов

3.4.1 Ключевые выводы по часам

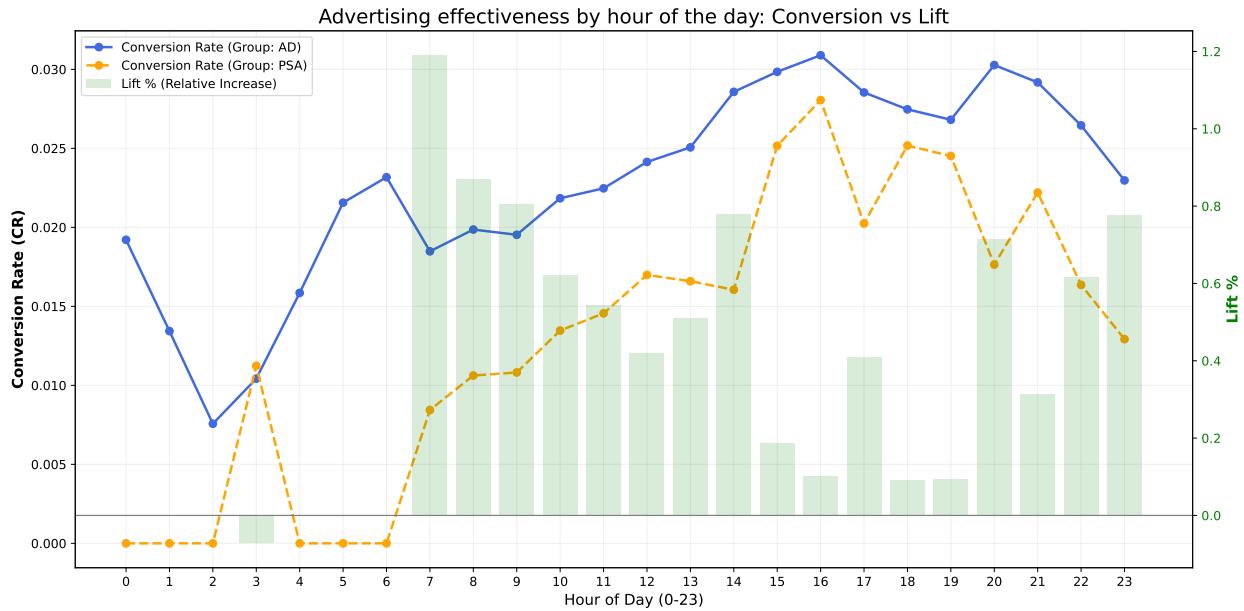


Рис. 2: График эффективности рекламы в зависимости от времени суток. Оранжевым отложенена конверсия в контрольной группе, синим - в тестовой. Зелеными столбиками отмечен прирост конверсии в процентах.

Наблюдение: Реклама демонстрирует наивысшую конверсию во второй половине дня 13:00–23:00. Это совпадает с известными паттернами поведения потребителей: люди более активны в конце рабочего дня и вечером.

3.5 Анализ зависимости от частоты показов (Frequency Capping)

3.5.1 Тренд конверсии по бакетам

Диапазон показов	CR	Кол-во юзеров	Примечание
1–10	0.33%	249 499	Низкое вовлечение
11–20	0.84%	123 334	Начало роста
21–50	2.92%	125 541	Ключевой сегмент
51–100	11.63%	44 149	Резкий скачок
101–150	17.91%	11 854	Пик эффективности
151–200	17.07%	4 506	Плато (насыщение)
201–300	15.20%	3 389	Снижение (выгорание)
301–500	15.70%	1 739	Редкие активные юзеры
500+	17.49%	566	Редкие выбросы

Таблица 2: Конверсия в зависимости от частоты показов (только группа AD)

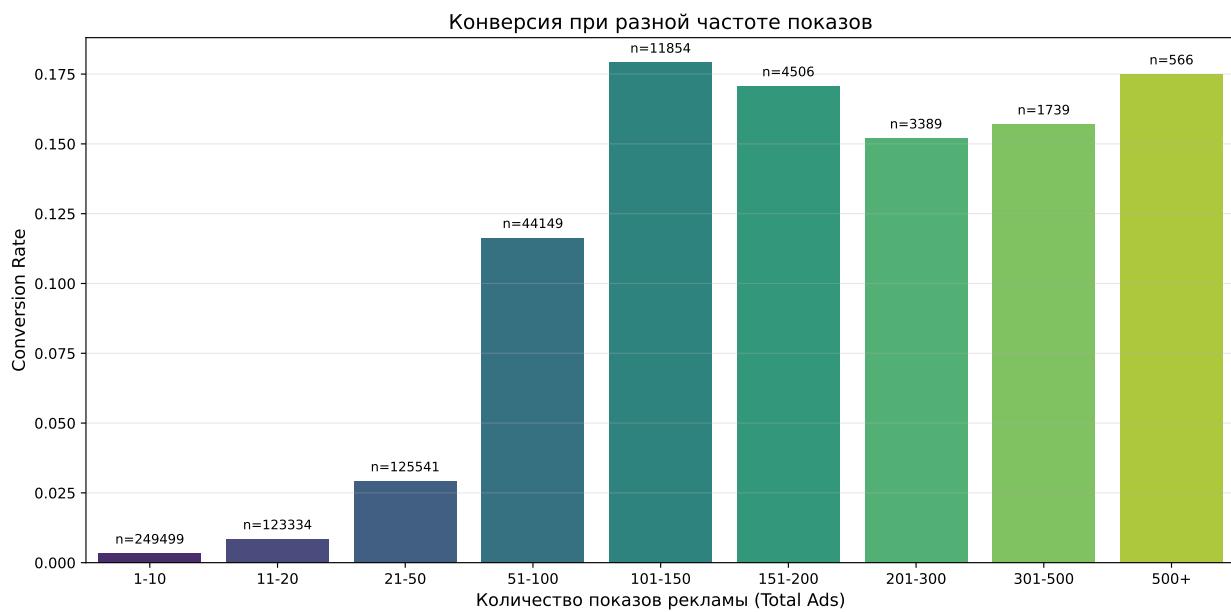


Рис. 3: Конверсия в зависимости от частоты показов

3.5.2 Интерпретация

Пользователи, которые изначально были заинтересованы в покупке, преследовались рекламой чаще (механизм ретаргетинга). Это типичный пример обратной причинности.

Точка насыщения: Реальный эффект частоты проявляется на диапазоне 21–50 показов, где конверсия растет с 1% до 3%. После 100 показов рост практически прекращается.

Бизнес-рекомендация: Frequency cap на уровне 50–100 показов в период будет оптимальен. Дальнейший рост числа показов приносит минимальный ROI при увеличении раздражения пользователей.

4 Результаты

4.1 Глобальные показатели эффективности

По итогам эксперимента рекламная кампания продемонстрировала высокую эффективность по сравнению с контрольной группой.

Метрика	Группа AD (Тест)	Группа PSA (Контроль)	Разница
Пользователи (n)	564 577	23 524	—
Конверсии	14 423	420	—
Conversion Rate	2.55%	1.79%	+0.76 p.p.

Таблица 3: Сводные результаты A/B теста

4.2 Статистическая значимость

Для проверки гипотезы о том, что наблюдаемый рост конверсии не является случайным выбросом, были применены два статистических критерия:

- Z-test для пропорций:
- Хи-квадрат Пирсона:

Вывод: Так как P-value многократно меньше уровня значимости $\alpha = 0.05$, мы отвергаем нулевую гипотезу о равенстве конверсий. Различие статистически значимо и носит системный характер. Вероятность получить такие результаты случайно близка к нулю.

4.3 Оценка Lift и доверительные интервалы

Ключевой метрикой для бизнеса является относительный прирост (Lift), который показывает, насколько реклама увеличила вероятность покупки.

$$\text{Lift} = \frac{2.55\% - 1.79\%}{1.79\%} \times 100\% = 43.09\% \quad (5)$$

Для оценки надежности этой метрики был применен метод Bootstrap (2000 итераций), который позволил построить 95% доверительный интервал для Lift:

- Средний Lift: 43.55%
- 95% Доверительный интервал (CI): [30.25%, 58.13%]

Интерпретация для бизнеса: С вероятностью 95% истинный эффект от рекламы лежит в диапазоне от +30% до +58%. Даже в пессимистичном сценарии (нижняя граница) реклама увеличивает продажи почти на треть.

4.4 Почасовая эффективность

Анализ конверсии в разрезе часов суток позволил выявить неравномерность эффекта рекламы:

- Пик эффективности (Prime Time): В период с 13:00 до 23:00 наблюдается максимальные значения конверсии как в тестовой, так и контрольной группах. В эти часы пользователи наиболее восприимчивы к рекламному предложению.
- Пик разницы: В период с 07:00 до 12:00 наблюдается максимальный разрыв между тестовой и контрольной группой. В эти часы эффект рекламы наиболее заметен.
- Спад: Вочные часы и ранним утром (00:00–06:00) конверсия в тестовой группе падает и практически сравнивается с контрольной. Lift в эти часы минимален или статистически незначим.

Рекомендация: Для оптимизации бюджета следует повысить ставки (bid adjustment) в вечерний слот (16:00–20:00) и снизить активность показов в ночное время, так как "органическая" конверсия ночью низка, а реклама не дает значимого прироста.

4.5 Анализ частоты (Frequency Capping)

Детальный анализ "хвоста" распределения показов (см. Таблицу 2) показал, что рост конверсии не бесконечен.

- Зона роста: До 50 показов конверсия растет.
- Зона плато: Максимальная конверсия (17.91%) достигается в сегменте 101–150 показов и дальше почти не меняется.

Это подтверждает гипотезу о насыщении: после 100-го контакта реклама перестает работать эффективно и начинает вызывать "баннерную слепоту" или раздражение.