

Презентация к проекту А/В-теста

✂ ЗАДАЧА:

Протестировать изменение расположения рекламы: для тест-группы секция рекламы должна будет располагаться над основной секцией.

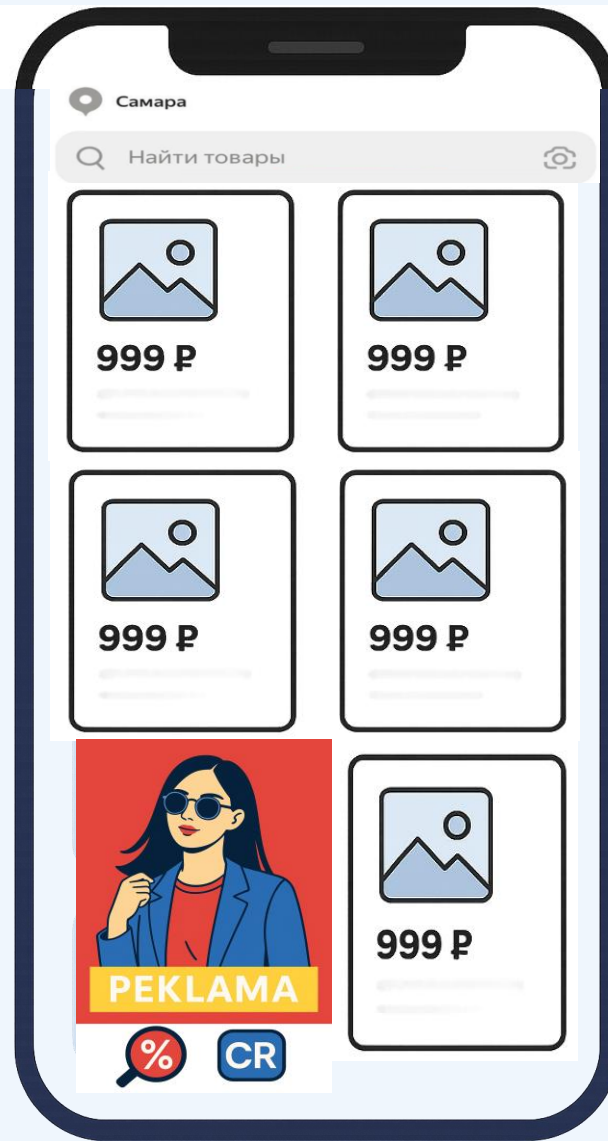
🎯 ЦЕЛЬ:

Проверить, влияет ли новое расположение рекламного блока в приложении маркетплейса на поведение пользователей.

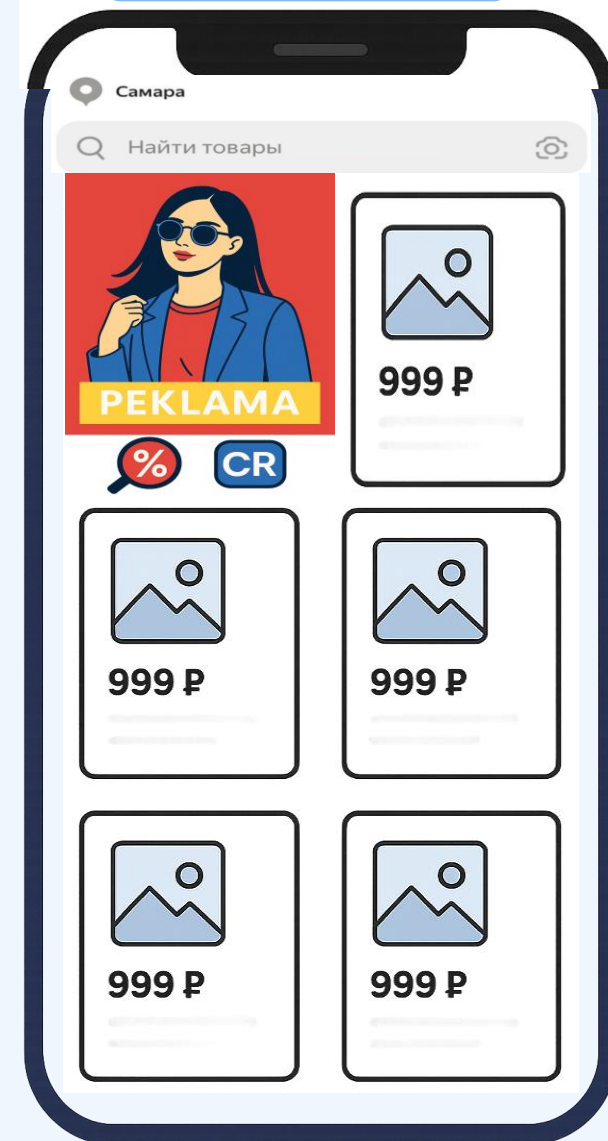
💡 ГИПОТЕЗА:

Новое расположение увеличивает вовлечённость и ключевые пользовательские метрики.

КОНТРОЛЬ



ТЕСТ



ДИЗАЙН ЭКСПЕРИМЕНТА И РАСЧЕТ MDE

ОГРАНИЧЕНИЯ И ПРЕДПОСЫЛКИ



20 %



CR + 2,5 %



1 неделя

✦ ОСНОВНЫЕ МЕТРИКИ, КОТОРЫЕ ХОТИМ ОТСЛЕЖИВАТЬ:



Конверсия в добавление товаров
(adds / clicks) (CR)



КЛЮЧЕВАЯ метрика



Конверсия в просмотр (CTR)
(клики / просмотры)



Дополнительная метрика



Конверсия в добавление от просмотра главной
(Conv_view_to_add)



Дополнительная метрика



Конверсия в заказ от захода в рекламу
(заказы / клики)



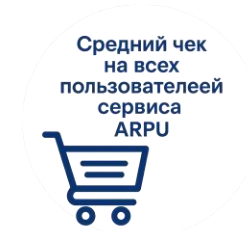
Нет данных ДО (заказов)



Конверсия в заказ от просмотра главной
↑ заказы / просмотры



Нет данных ДО (заказов)



Средний чек на всех пользователей сервиса
ARPU



Нет данных ДО (сумм)



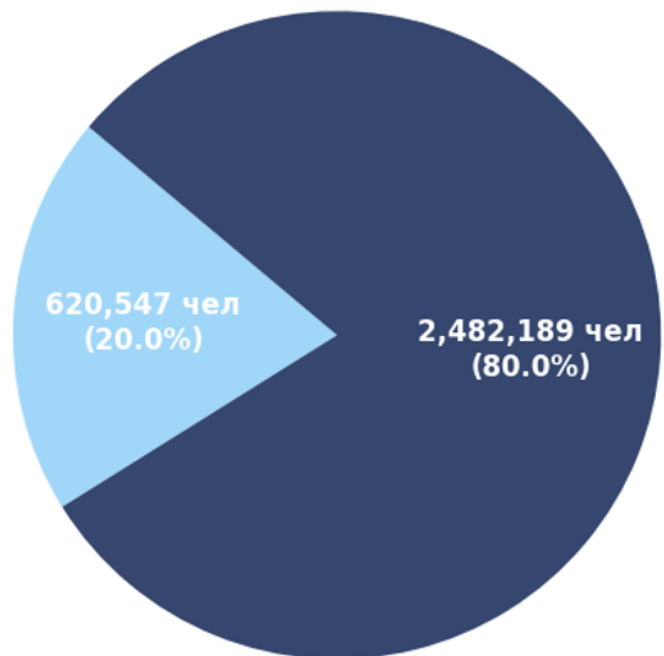
Выручка на покупателя
ARPPU
₽



Нет данных ДО (сумм)

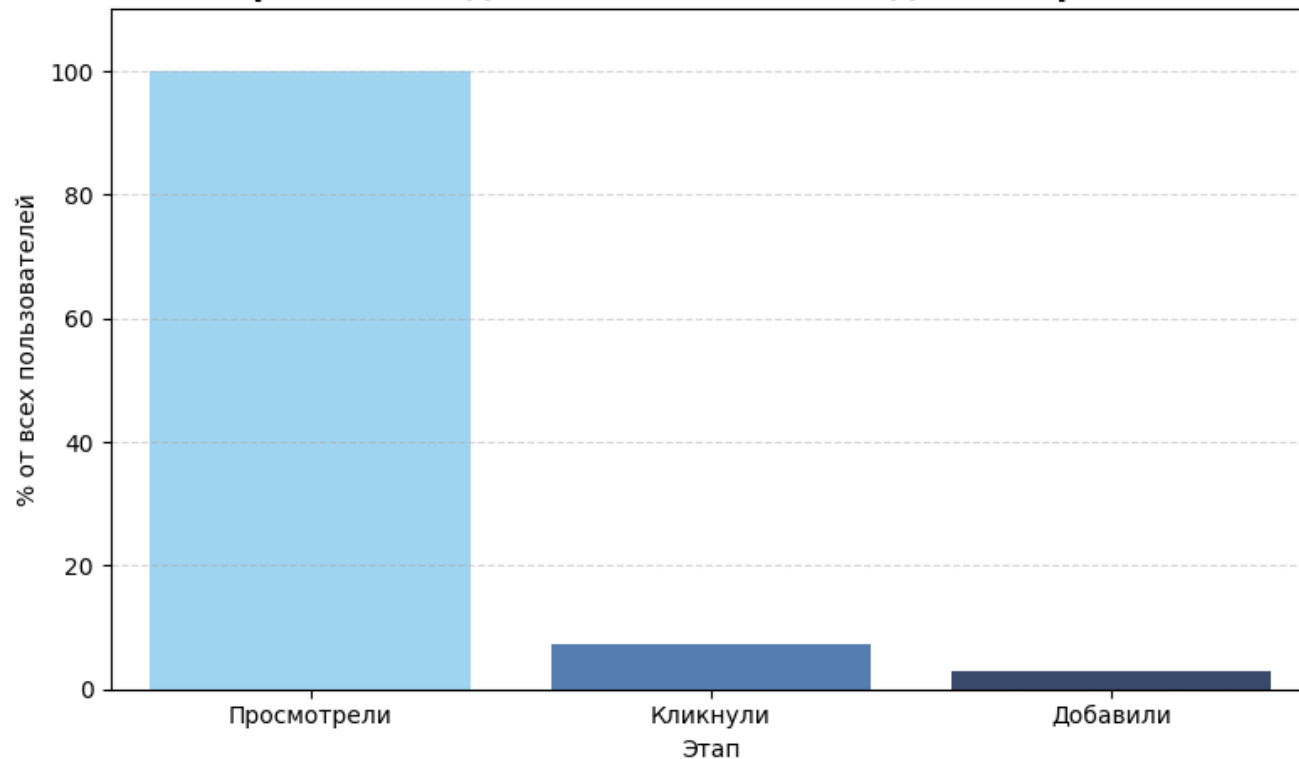
Изучение пользовательского поведения до эксперимента

Распределение пользователей до эксперимента
(по участию в A/B-тесте)



Уникальных пользователей: 3 102 736

Воронка поведения пользователей до эксперимента



Этап % пользователей

0	Просмотрели (views)	100.00
1	Кликнули (clicks)	7.36
2	Добавили (adds)	2.91



Ключевая метрика А/В - теста

Условия теста:

- Объём выборки:

В совокупности 620 547 пользователей
(по 310 273 в тест и контроль)

- Параметры:

$\alpha = 0.05$,

$1-\beta = 0.80$,

длительность — 1 неделя

КЛЮЧЕВАЯ МЕТРИКА

→ **CR – Конверсия в заказ**

Финальное целевое действие:
пользователь оформил заказ
после взаимодействия с
рекламой

ДОПОЛНИТЕЛЬНЫЕ МЕТРИКИ

CTR

- ◆ Кликабельность рекламного блока
- ◆ Клики/Показы рекламы отражает вовлеченность

Просмотр
рекламы

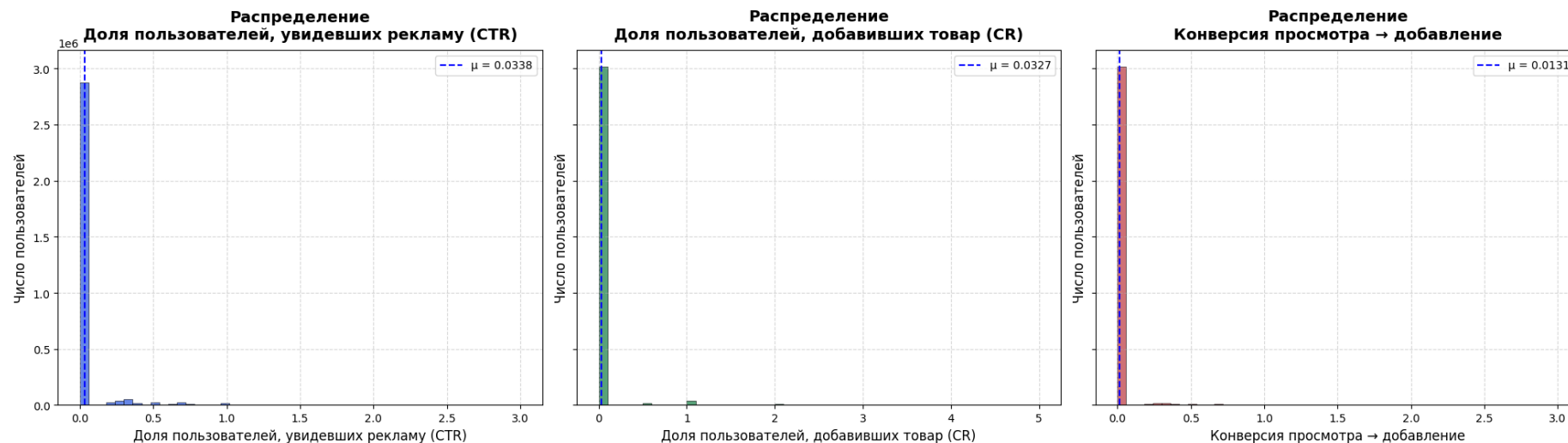
Оформление
заказа

ЗАЧЕМ НУЖНЫ ДЕТАЛИЗАЦИЯ

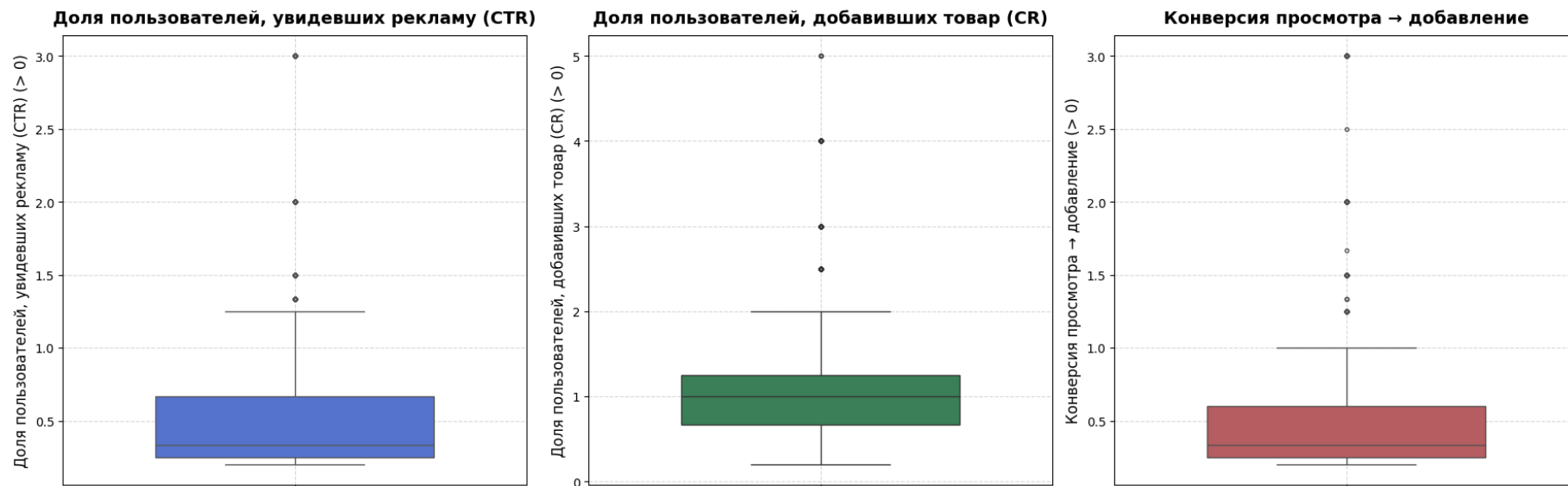
- ◆ Ключевая метрика отвечает на вопрос, помогает ли новое размещение довести до покупки
- ◆ Дополнительные дают понимание, как именно

В сведениях до эксперимента (в файле: data_raw.csv):

Гистограммы распределений по основным метрикам (нулевой период)



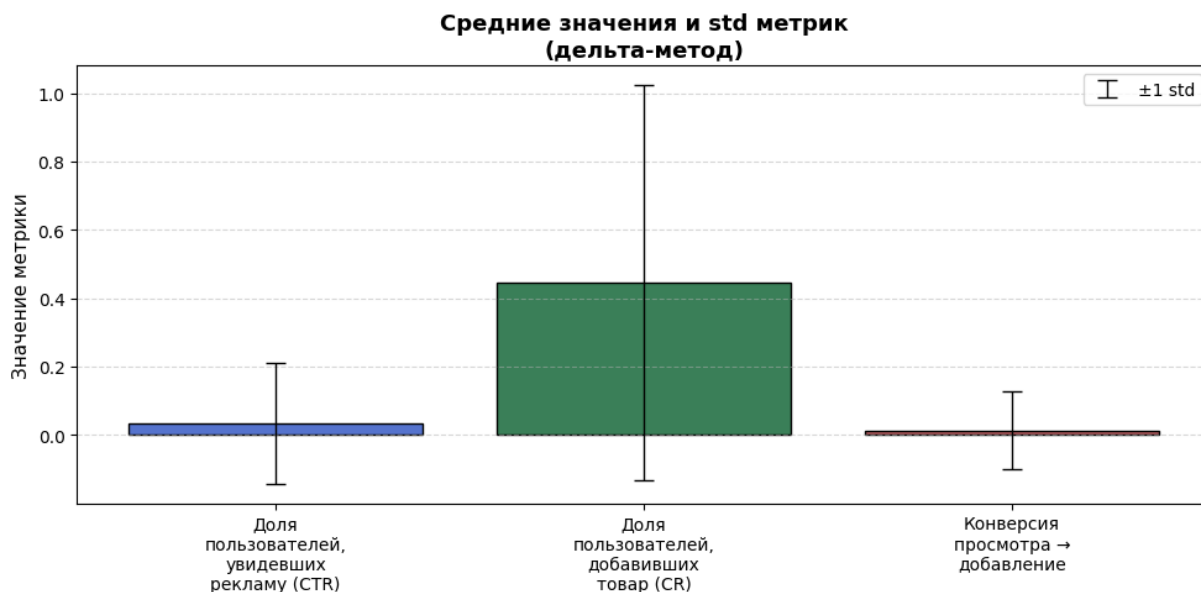
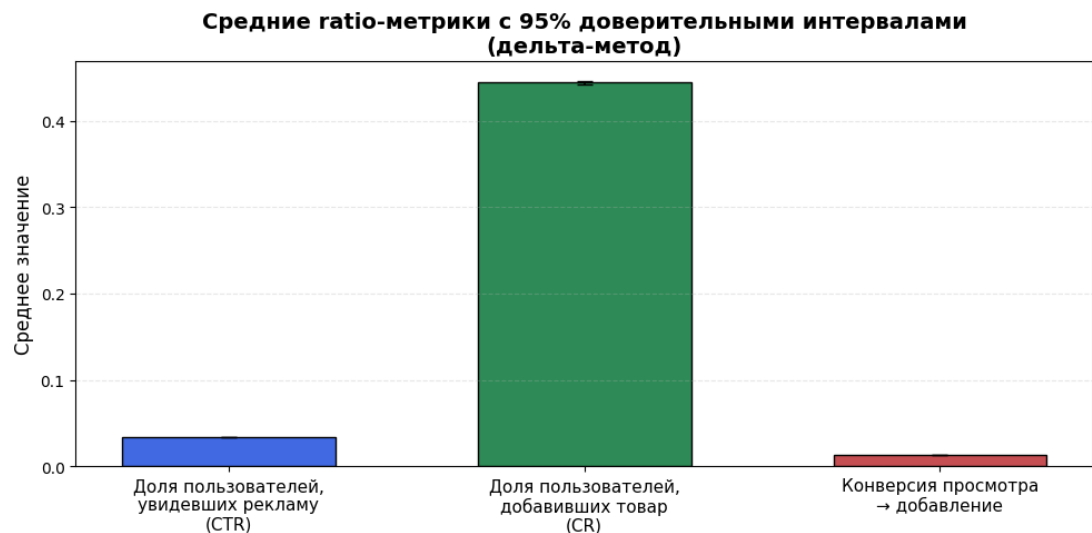
Boxplot распределений метрик для активных пользователей



Распределения метрик сильно скошены и содержат выбросы.

Большинство пользователей не совершают целевых действий — необходим дельта-метод и стратификация

Сравнение средних значений метрик с доверительными интервалами и стандартными отклонениями (дельта-метод)



🇷🇺 95% доверительные интервалы для ratio-метрик:

	Метрика	mean	se	ci_lower	ci_upper
0	Доля пользователей, увидевших рекламу (CTR)	0.033769	0.000100	0.033574	0.033964
1	Доля пользователей, добавивших товар (CR)	0.444304	0.001210	0.441932	0.446676
2	Конверсия просмотра → добавление	0.013081	0.000065	0.012953	0.013208

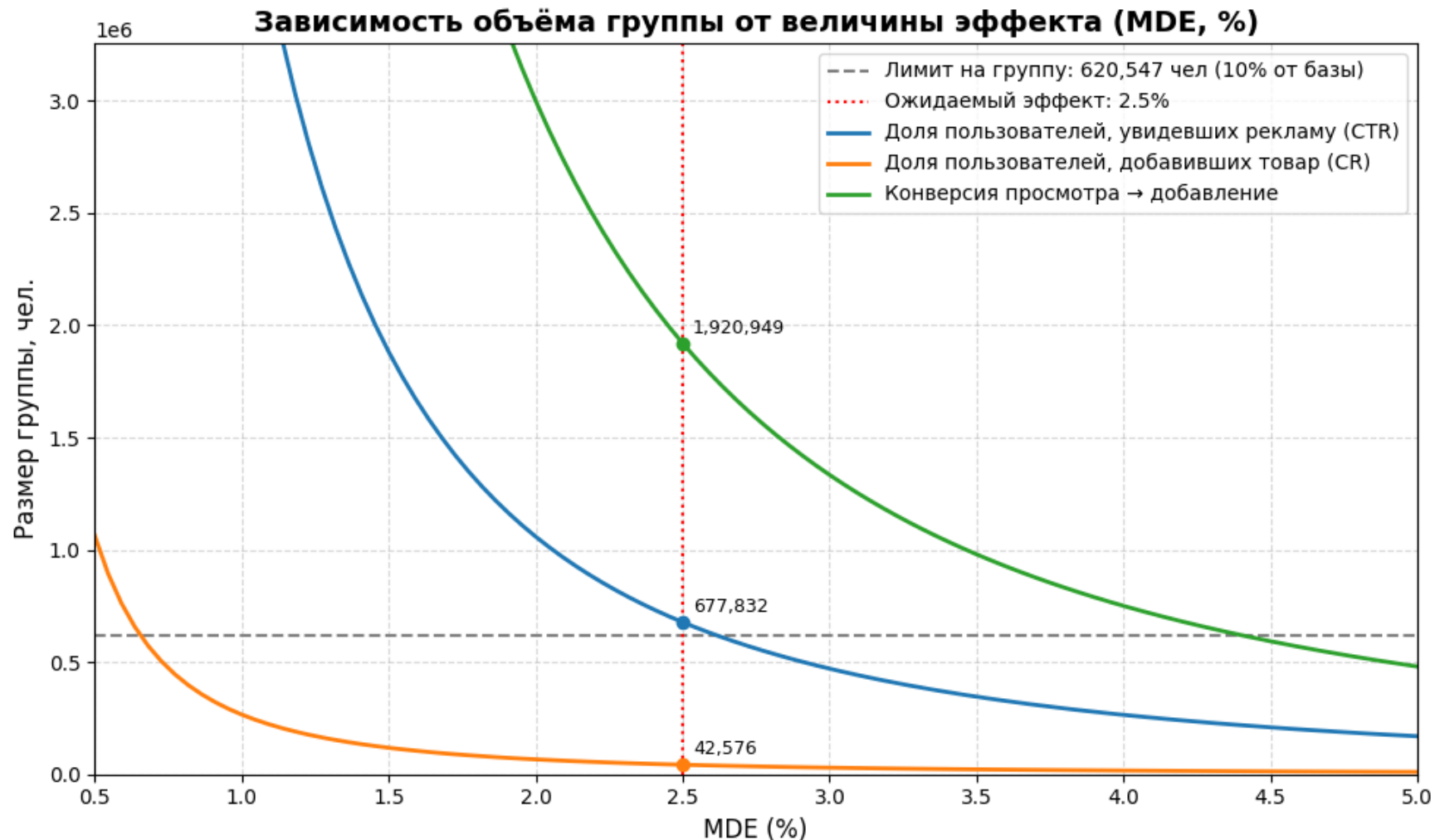
Вывод: Несмотря на значительные различия в вариативности ratio-метрик, 95% доверительные интервалы остаются узкими благодаря большому объёму выборки. Это подтверждает:

- ◆ **Обоснованность дельта-метода** для оценки дисперсии дробных метрик,
- ◆ **Надёжность средних значений** при сравнении контрольной и тестовой групп.

Чувствительность теста: Зависимость объема выборки от эффекта (MDE)

Интерпретация:

- Зафиксировать эффект +2.5% по **CTR** и **Conv_view_to_add** (Конверсия просмотров в добавление) невозможно в рамках текущего лимита (20% пользователей);
- Только по метрике **CR** объём выборки достаточен для выявления такого эффекта.



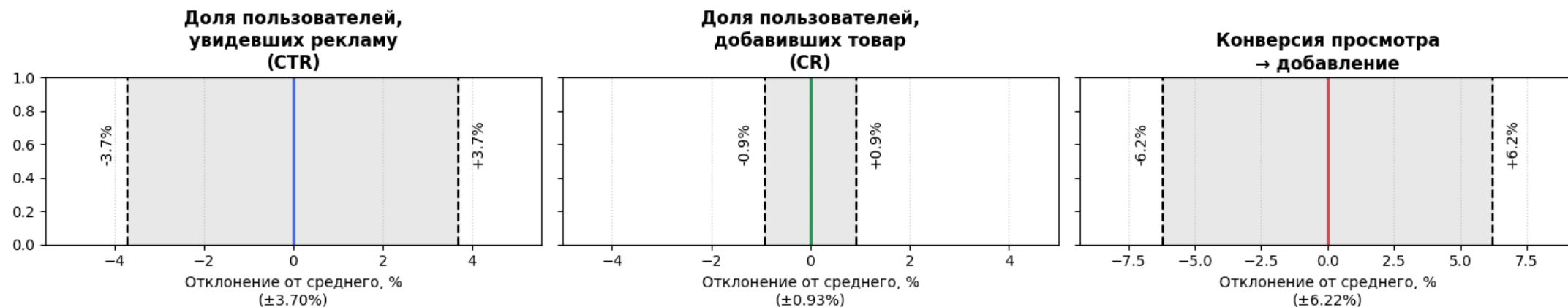
Вывод:

Эффект +2.5% реалистично протестировать только по метрике **CR**.
Для остальных метрик требуется либо **большая выборка**, либо **дольше тест**.

Минимальный различимый эффект (MDE, %) при текущем объеме выборки.

Диапазон чувствительности A/B-теста ($\pm MDE$)

Диапазон чувствительности A/B-теста
(заштрихована зона $|\Delta| < MDE$)



📌 Заключение:

При охвате 20% можно протестировать гипотезу только по метрике **CR**.
Остальные метрики подходят лишь для **сильных эффектов** или дополнительного анализа.

Примечание:

Заштрихованная область $|\Delta| < MDE$ показывает, какие изменения тест **не сможет** отличить от случайного шума при текущем размере выборки. Изменения, выходящие за пределы $\pm MDE$ (линии на 100 % $\pm MDE$), будут статистически различимы.

Расчёт минимально различимого эффекта (MDE) при заданном размере групп (по 10%):

	Метрика	Код метрики	Среднее значение	Дисперсия (дельта)	Ист. число пользователей	MDE (абс.)	MDE (%)
0	Доля пользователей, увидевших рекламу (CTR)	CTR	0.033769	0.030775	3102735	0.00125	3.70
1	Доля пользователей, добавивших товар (CR)	CR	0.444304	0.334631	228491	0.00411	0.93
2	Конверсия просмотра → добавление	Conv_view_to_add	0.013081	0.013086	3102735	0.00081	6.22

ОБЩИЙ ВЫВОД (ЧАСТЬ 1)

Методологическая часть подтверждает: тест корректно спроектирован и способен зафиксировать ожидаемый эффект по ключевой метрике CR.



Охват

620 557 пользователей
(20 % аудитории)



Рассчитанный MDE

$\pm 0,93 \%$



Ключевая мерка

CR – конверсия в
добавление после клика



Вывод

Эффект фиксируется
при выборке $\geq 42\,576$
пользователей



Целевой эффект

+ 2,5 %

ГОТОВНОСТЬ К АНАЛИЗУ РЕЗУЛЬТАТОВ:

Методологическая часть завершена.



Переходим к **ЧАСТИ 2** – оценке фактических результатов эксперимента и проверке гипотезы.

Баланс групп: корректность А/В-распределения

Включает уникальных пользователей, показы, добавления, заказы и выручку

Сводная таблица абсолютных значений (control / test / Δ)

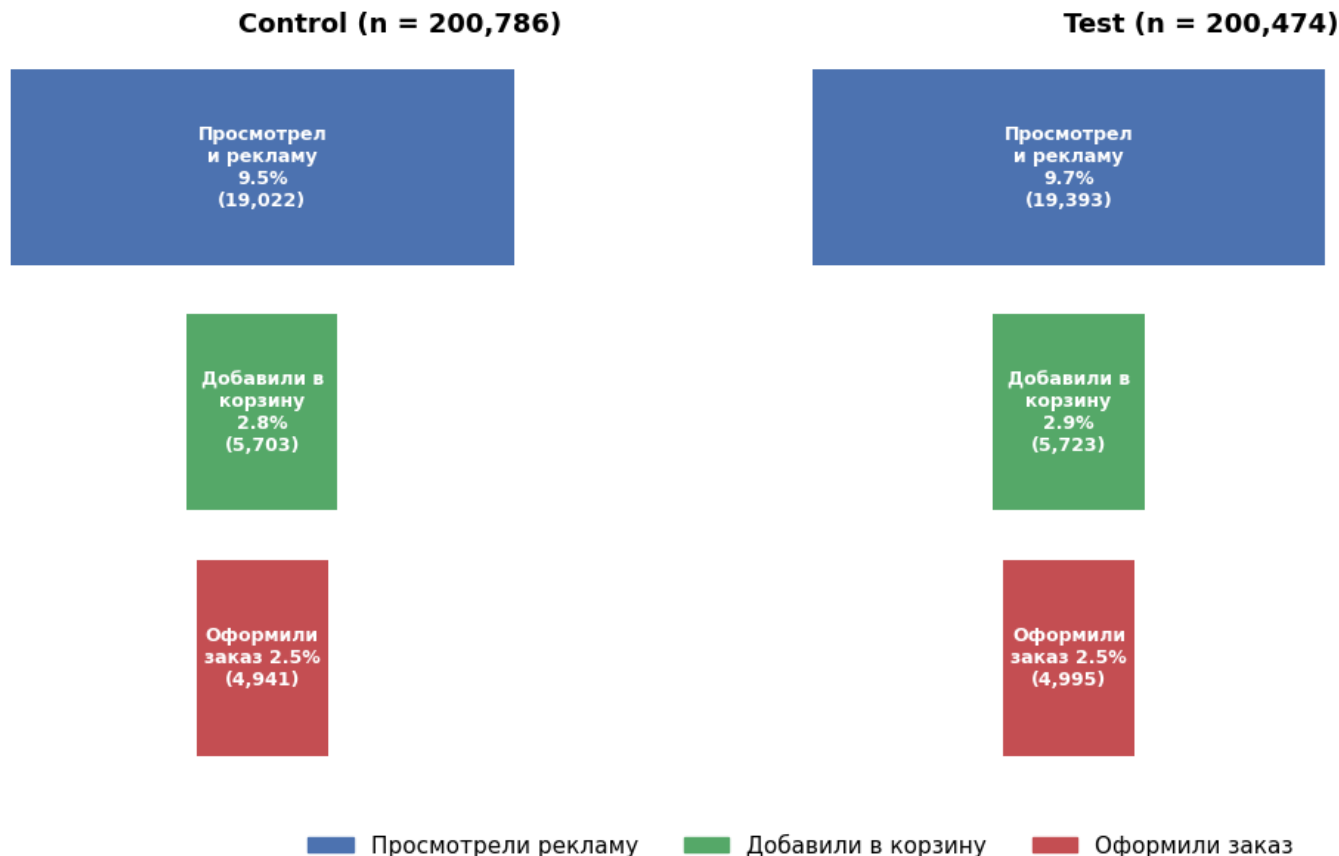
ab_group	Показатель	control	test	Разница (test-control)
0	Уникальные_пользователи	200,786	200,474	-312
1	Дней_с_просмотрами	25,389	25,852	+463
2	Всего_показов	43,469	43,805	+336
3	Дней_с_добавлениями	7,060	7,085	+25
4	Всего_добавлений	15,799	16,337	+538
5	Стоимость_добавлений	2,462,420	2,526,393	+63,973
6	Дней_с_заказами	6,079	6,171	+92
7	Всего_заказов	5,579	5,881	+302
8	Выручка_из_заказов	1,589,899	1,687,895	+97,996



Вывод:

Обе группы имеют почти равное количество пользователей (разница < 0.2%).
Распределение сбалансировано — можно корректно сравнивать метрики без дополнительных корректировок.

Воронки пользовательских действий (контроль vs тест)



Интерпретация:

- ◆ Просмотры рекламы в тестовой группе **незначительно выросли**: 9.5% → 9.7% (почти на том же уровне).
- ◆ Добавления в корзину **увеличились**: +0.1 п.п. (2.8% → 2.9%) — пусть и слабо, но в сторону улучшения.
- ◆ Оформление заказов осталось на **прежнем уровне**: 2.5% в обеих группах — это говорит об отсутствии сдвига на финальном этапе воронки.

Вывод:

- ◆ Поведение пользователей **на ранних этапах** воронки (просмотры → добавления) слегка улучшилось.
- ◆ **Финальная конверсия в заказ** — без изменений, что может быть признаком **недостаточного влияния размещения рекламы** на реальные продажи.
- ◆ Для принятия решения о внедрении стоит учитывать **влияние на глубинные метрики (ARPU, CR)** и **результаты статистического теста**.

Сравнение метрик по группам

Группы сбалансированы — можно переходить к анализу статистической значимости.

Сравнение ключевых метрик Умеренные улучшения в тестовой группе.					
Метрика	Контроль	Тест	Δ (pp / абс.)	Δ (%)	Комментарий
Доля пользователей с показами рекламы	9.47 %	9.67 %	+0.20 pp	+2.1 %	Слегка вырос охват – почти на 2 % относительно контроля.
Среднее число показов рекламы на пользователя	0.2165	0.2185	+0.0020	+0.9 %	Количество показов на пользователя практически не изменилось.
Доля пользователей, добавивших товар	2.84 %	2.85 %	+0.01 pp	+0.4 %	Роста доли «добавлений в корзину» почти нет.
Среднее число добавлений из рекламы на пользователя	0.0787	0.0815	+0.0028	+3.6 %	Пользователи чуть активнее добавляют товары после рекламы.
Средняя сумма добавлений из рекламы	12.2639 (валюта)	12.6021 (валюта)	+0.3382	+2.8 %	Средний чек «добавлений» вырос на ~0.34.
Доля пользователей, оформивших заказ	2.46 %	2.49 %	+0.03 pp	+1.2 %	Незначительный рост конверсии в заказ.
Среднее число заказов на пользователя	0.0278	0.0293	+0.0015	+5.4 %	Количество заказов на пользователя выросло заметнее.
Средняя сумма заказов на пользователя	7.9184 (валюта)	8.4195 (валюта)	+0.5011	+6.3 %	Выручка на пользователя выросла примерно на 0.50.

Основные выводы

- Небольшой, но стабильный рост** во всех метриках тестовой группы по сравнению с контролем.
- Максимальный относительный прирост** наблюдается в среднем числе заказов на пользователя (+5.4 %) и в среднем доходе на пользователя (+6.3 %). Это ключевые коммерческие метрики, наиболее релевантные для оценки успеха эксперимента.
- Малые изменения** в долях пользователей с показами рекламы и конверсии «просмотр → добавление» находятся в пределах 0–3 %, что может быть следствием статистического шума.
- Рекомендация:** следующий шаг — рассчитать дельты, доверительные интервалы и р-значения для каждой метрики, чтобы оценить, какие из этих сдвигов действительно статистически значимы.

Сравнение средних ключевых метрик в эксперименте: Test vs Control

**CTR**

(показы →
клики)

$\Delta \approx 0,0020$ п.п.
(+ 2,11 %)

- Незначительное улучшение, но не превысил MDE 2,5%

**CR**

(клики →
добавления)

$\Delta \approx 0,0001$ п.п.
(+ 0,35 %)

- Эффект отсутствует

**Конверсия из
показа в
добавление**

$\Delta \approx 0,0028$ п.п.
(+ 3,56 %)

- Незначительный рост, превышает MDE 2,5 %

**Выручка на
пользователя**

$\Delta + 0,5011$ п.п.
(примерно + 6,3 %)

- Наибольший рост



*Выручка на пользователя в эксперименте показала наибольший рост (+6.3%).
Остальные метрики требуют статистической проверки для подтверждения эффекта.*

ФИНАЛЬНОЕ СРАВНЕНИЕ: разность между группами и границы чувствительности (MDE)

95% CI, MDE (%)

**CTR**

ДОЛЯ ПОЛЬЗОВАТЕЛЕЙ
УВИДЕВШИХ РЕКЛАМУ
MDE (абс.) $\approx +0.00260$ п.п.

2.75 %

Контроль

Тест



Статистически значимо, но меньше
MDE (зато $> 2.5\%$)



Конверсия в
добавление
товаров
(adds / clicks) (CR)

CR

ДОЛЯ ПОЛЬЗОВАТЕЛЕЙ
ДОБАВИВШИХ ТОВАР
MDE (абс.) $\approx +0.00147$ п.п.

5.18 %

Контроль

Тест



доверительные интервалы полностью
перекрываются. По ключевой метрике
CR нет никакого прироста.

Основная метрика CR не выросла: тест показал
практически нулевое изменение, и его доверительный
интервал полностью совпадает с контрольным.

**Conv_view_to_add**

ДОБАВЛЕНИЕ ОТ ПРОСМОТРА
MDE (абс.) ≈ -0.00738 п.п.

9,37 %

Контроль

Тест



CI существенно перекрываются, то
есть изменение статистически
незначимо.

**Revenue_per_user**

ВЫРУЧКА НА ПОЛЬЗОВАТЕЛЯ
MDE (абс.) $\approx +0.72932$ п.п.

9,21 %

Контроль

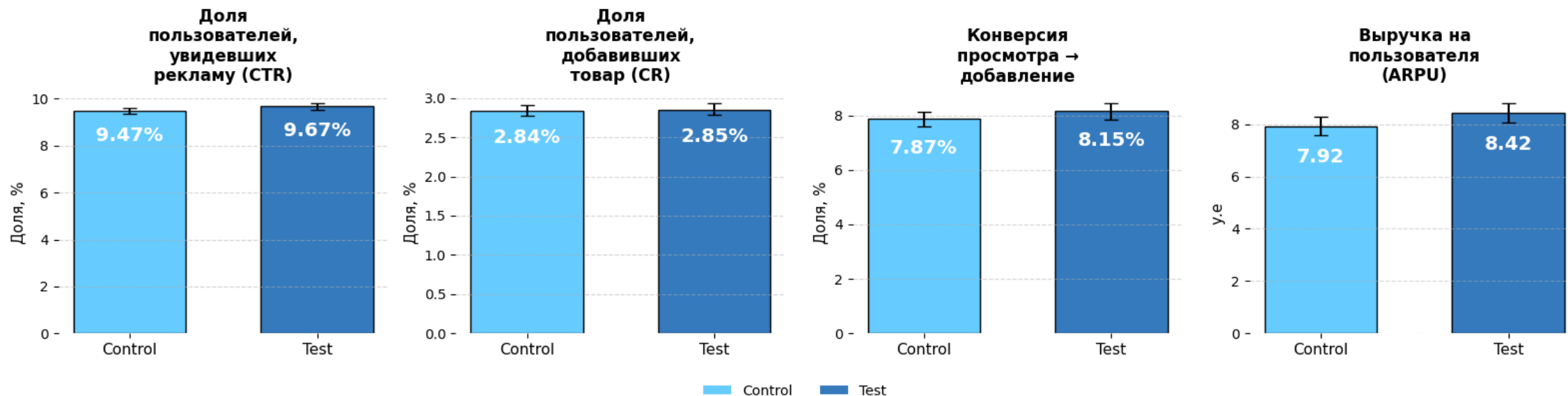
Тест



Доверительные интервалы пересекаются
в диапазоне 8.048 ₹ - 8.268 ₹,
поэтому по CI-разнице однозначной
статистической значимости тоже нет,
хотя тренд на увеличение выручки
налицо.

ФИНАЛЬНОЕ СРАВНЕНИЕ: разность между группами и границы чувствительности (MDE)

Сравнение метрик между группами с 95% CI



Вывод:

- Ключевая метрика CR практически не изменилась — доверительные интервалы полностью перекрываются → эффект отсутствует.
- CTR и ARPU показали лёгкий рост, однако статистическая значимость не достигнута — CI перекрываются.
- Conv_view_to_add незначительно снизилась, но также без статистической значимости.

Заключение:

Изменение расположения рекламы **не привело к значимому улучшению CR**. Остальные метрики демонстрируют лишь слабые сдвиги. Для уверенных выводов рекомендуется повторный А/В-тест с увеличенной выборкой или альтернативным вариантом интерфейса.

Неожиданные эффекты эксперимента

Несмотря на слабый эффект по основным метрикам, выявлены положительные изменения в поведении пользователей и выручке

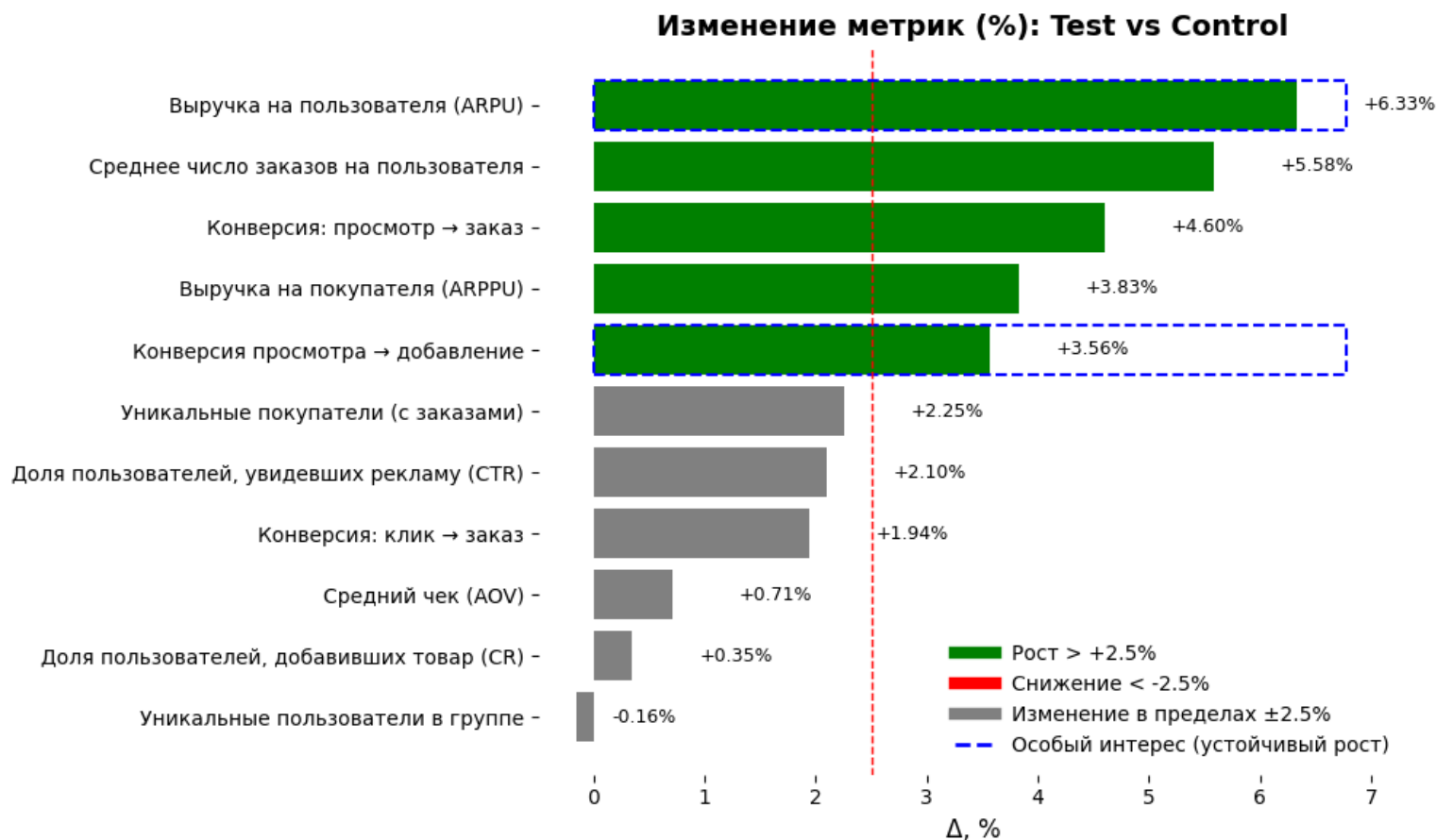


Порог чувствительности установлен на уровне $\pm 2.5\%$. Только метрики, превысившие этот порог, рассматриваются как практически значимые.

Метрики сформированы на основе пользовательских действий за период эксперимента, рассчитаны на одного пользователя или покупателя.

Вывод:

Несмотря на слабый эффект по основной метрике (CR), эксперимент выявил **устойчивые положительные изменения в выручке и микро поведенческих показателях**. Это подтверждает гипотезу о пользе нового расположения рекламы — особенно в части доходности.



Поведение пользователей до и после эксперимента



ПРОСМОТРЫ

ДО - 310 273 чел.
ПОСЛЕ - 200 474 чел.



ДОБАВЛЕНИЯ

ДО - 310 273 чел.
ПОСЛЕ - 200 474 чел.



Анализ изменения пользовательской активности (До → После, независимые выборки)

	Метрика	n_pre	n_post	Среднее ДО	Среднее ПОСЛЕ	Δ	Изменение (%)	95% CI низ	95% CI верх	Стат. значимо	Направление	Комментарий
0	Просмотры	310,273	200,474	2.6677	0.1290	-2.5387	-95.17%	-2.5434	-2.5341	✓	↓ Падение	Существенное падение
1	Добавления	310,273	200,474	0.0478	0.0815	+0.0337	+70.34%	0.0306	0.0367	✓	↑ Рост	Существенный рост

✂ Методология:

◆ Использовался дельта-метод для анализа **независимых выборок**:

ДО — случайная доля пользователей (10%), ПОСЛЕ — вся тест-группа.

◆ Сравнение основано на **абсолютных действиях пользователей** — количество просмотров и добавлений, а не на долях или ratio-метриках.

ВЛИЯНИЕ МОГЛИ ОКАЗАТЬ:

сезонность

маркетинговые кампании

внешние события

ошибки измерений



Что мы на самом деле измеряем?

Изменения поведения ДО и после \neq Эффект эксперимента



Сравнение средней активности по test-когорте ДО и ПОСЛЕ



	Число клиентов	До	После	Разница
Просмотры	49670.0	2.8296	0.1708	-2.6588
Добавления	49670.0	0.0596	0.0902	0.0305



Да, поведение пользователей изменилось во всех группах, и эти изменения статистически значимы



Но, внутри самого эксперимента (test и control) различий не наблюдается – метрики не «прокрасились».

===== ВЫВОД =====

Мы фиксируем влияние внешних факторов (например: сезонность, промо-акции и пр.), а не результат нового расположения рекламы.



Только сравнение между группами в одном и том же периоде дает нам основание делать вывод о влиянии теста.



Корректное сравнение — только между группами test и control в один и тот же период (с учётом MDE, CI, p-value и т.д.)

ФИНАЛЬНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ



Ключевая метрика – CR

(доля пользователей, добавивших товар после показа рекламы)



Эффект: + 0.02 п.п.



Не достигнута статистическая значимость



Не превышен порог MDE (2.5 %)

Гипотеза не подтверждена – текущее изменение нельзя считать значимым

Дополнительные сигналы:

Выручка на пользователя



ARPU



+ 6,3 %



рост выручки с пользователя

Выручка на покупателя



ARPPU



+ 3,8 %



рост выручки на покупателя

Конверсия просмотра в добавление



Conv_view_to_add



+ 2,6 %



улучшение глубинных конверсий

Конверсия просмотра в заказ



Conv_view_to_order



+ 4,6 %



движение к заказу стало эффективней

ФИНАЛЬНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ

РЕКОМЕНДАЦИИ МЕНЕДЖЕРУ:



Не внедрять изменения немедленно:



хотя выручка выросла статистически значимо, эффект не превышает MDE — возможно, это шум.



Провести повторный эксперимент:



- на большей выборке,
- с более длительным горизонтом,
- с анализом по сегментам (новые / лояльные пользователи, источники трафика).



Углубить анализ по выручке и чеку

