Вкладка 1

**A/B-тест:   
Изменение размещения рекламного блока в приложении маркетплейса**

Выполнил: Матвеев В.Е.

# **Описание проекта**

**Цель проекта** - проверка гипотезы, что изменение размещения рекламы на странице маркетплейса влияет на взаимодействие пользователей с рекламой.

Компания является маркетплейсом товаров. В приложении есть рекламные блоки, с помощью которых продавцы могут продвигать свои товары. Также через них можно информировать об акциях и скидочных днях, например, о Черной пятнице.

Сейчас первый рекламный блок расположен после 3 линий товаров бесконечной ленты (ленту товаров можно скроллить и рекламные блоки будут появляться среди товаров с какой-то периодичностью). В данном эксперименте мы будем рассматривать только ПЕРВЫЙ рекламный блок.

Сейчас первый рекламный блок размещается в ленте товаров, менеджер хочет подвинуть рекламный блок наверх и сравнить результаты.

Примеры размещения рекламы приведены ниже. При клике на рекламный блок открывается подборка товаров.

| Контроль | Тест |
| --- | --- |
|  |  |

**Задача**: Необходимо протестировать изменение расположения рекламы. Для тест-группы секция рекламы должна будет располагаться над основной.

Проверка гипотезы осуществляется через проведение **А/Б теста**.

**Задачи, решаемые при проведении А/Б теста:**

* Определение дизайна А/Б теста (параметров А/Б теста)
  + Формулирование гипотезы
  + Определение ключевых и вспомогательных метрик
  + Определение MDE
  + Определение минимальной аудитории для проведения А/Б теста
  + Определение времени теста
* Анализ и обработка исходных данных (проверка на дубли, пропуски)
* Анализ пользовательского поведения до проведения теста (анализ ключевых и вспомогательных метрик)
* Анализ и обработка итоговых данных А/Б теста
  + Анализ и обработка итоговых данных А/Б теста (проверка на дубли, пропуски)
  + Проверка на пересечение клиентов в контрольной и тестовой группах
* Анализ метрик в контрольной и тестовой группах и определение статистической значимости полученных результатов (расчет конверсий ведется по клиентам и дням)
  + Расчет ключевых и вспомогательных метрик
  + Расчёт uplift и статистической значимости (z-test, t-test)
  + Анализ изменения поведения пользователей при проведении теста

**Используемые инструменты:** Расчеты и визуализация проводились в notebook Google Collab (на языке Python).

Ссылка на Блокнот: [A/B-test](https://colab.research.google.com/drive/1SnnuJNrFD0DEVHcmIAp9q56i0xowOe_N?usp=sharing)

**Используемые библиотеки:** pandas, numpy, plotly.express, seaborn, matplotlib.pyplot, scipy.stats.

**Используемые данные:** [Пользовательские данные до теста](https://drive.google.com/file/d/1IoD_SoqL7gOHLWNbajn2xFENwu2XnHsy/view?usp=drive_link), [результат теста](https://docs.google.com/spreadsheets/d/1KRIuffR-x2GxSVX_8dkCnbDVj9bXONcV/edit?usp=drive_link&ouid=112984583747630199949&rtpof=true&sd=true)

# **Дизайн A/B-теста**

***Изменение размещения рекламного блока в приложении маркетплейса***

### **1. Цель эксперимента**

[[1]](#footnote-0)Проверить гипотезу: размещение первого рекламного блока выше (над лентой товаров) влияет на поведение пользователей, в том числе на кликабельность рекламы и конверсии в целевые действия (добавление в корзину).

### **2. Метрики**

#### **Ключевая метрика:**

* [[2]](#footnote-1)Конверсия в целевое действие **(добавление в корзину):**CR клик → добавление = Кол-во добавлений в корзину / Кол-во показов рекламы

#### **Вспомогательные метрики:**

* Кликабельность секции рекламы (CTR) = количество кликов / количество просмотров главной страницы
* Конверсия в заказ от просмотра главной страницы
* [[3]](#footnote-2)CR добавление → заказ - конверсия кол-ва заказов к кол-ву созданных корзин
* Процент выкупа корзины - сумма заказа к сумме товаров в корзине
* Добавление в корзину/ Показ рекламы - доля добавлений товаров в корзину после просмотра рекламы
* Средний чек товаров в корзине - сумма товаров в корзине, которые попали в корзину после просмотра рекламы, к кол-ву корзин
* Заказ / Добавление в корзину - доля заказов к количеству корзин
* Средний чек - сумма товаров в заказе к кол-ву заказов
* Кол-во Заказов на клиента в день
* Кол-во Заказов на клиента
* ARPU — выручка на клиента

### **3. Единица наблюдения**

* **Пользователь в день** — каждый пользователь может быть в выборке несколько раз, но анализ проводится на уровне уникального пользователя за день.

### **4. Минимально обнаруживаемый эффект (MDE)**

* Ожидаемый эффект рассчитывается на основе исходных данных, как **%** к ключевой метрике.

### **5. Параметры для расчета размера выборки**

* Уровень значимости (α): 0.05
* Мощность теста (1 - β): 0.8 (80%)
* MDE будет определена на основе анализа файла data\_raw.csv
* Базовая конверсия **(добавление в корзину)** будет определена на основе анализа файла data\_raw.csv

### **6. Дизайн теста**

* Тип эксперимента: **Randomized Controlled Trial (RCT)**
* Сплит: **50/50** (контроль / тест)
* Группа A (контроль): реклама после 3 линий товаров
* Группа B (тест): реклама над лентой товаров
* Продолжительность: **1 неделя**

### **7. Условия успеха**

* Тест считается успешным, если ключевая метрика (Конверсия в целевое действие **(добавление в корзину)**) в тестовой группе **значимо увеличивается** **или уменьшается** по сравнению с контрольной группой (p-value < 0.05).

# **Анализ структуры исходных данных и пользовательских действий**

### **1. Структура данных**

Исходные данные содержат 27 756 048 строк и 4 столбца. Дубляжей и пропусков нет.

### **2. Анализ пользовательской активности**

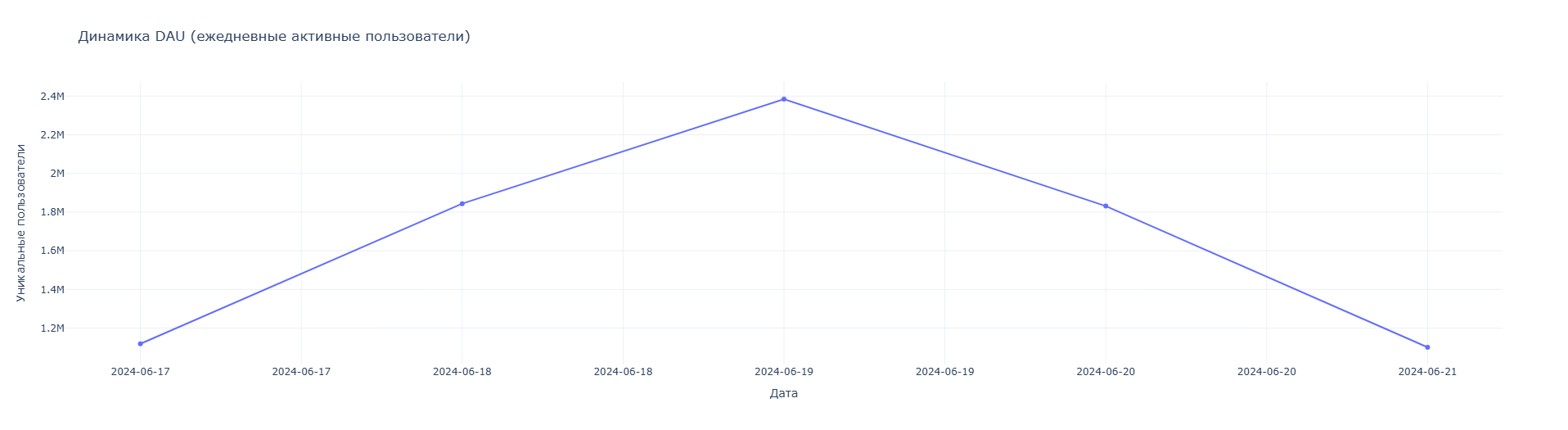
Основные характеристики пользователей представлены ниже:

#### **Кол-во уникальных клиентов и DAU**

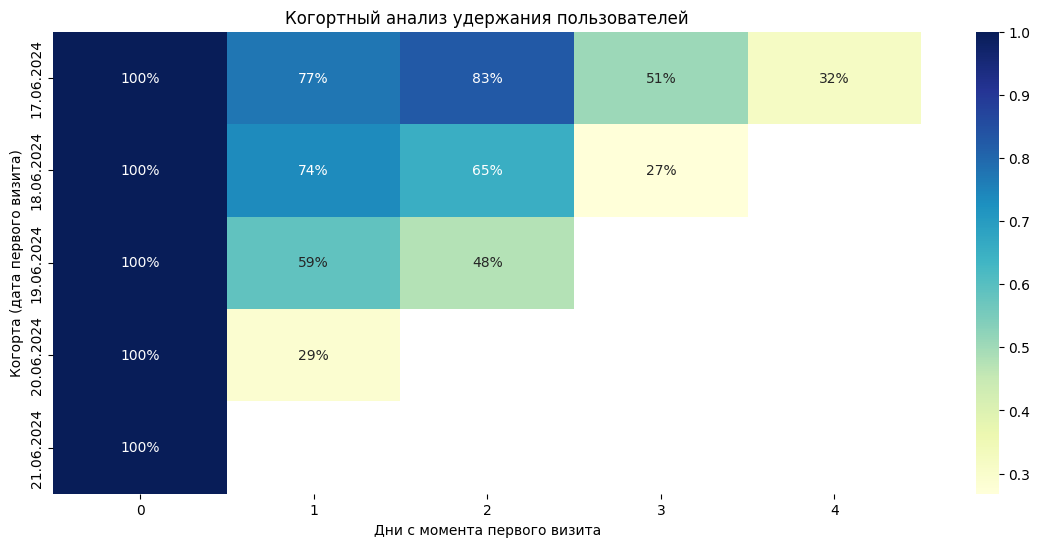
Количество уникальных клиентов: 3 102 736

Среднее DAU: 1 655 725.8

Медианное DAU: 1 831 314.0



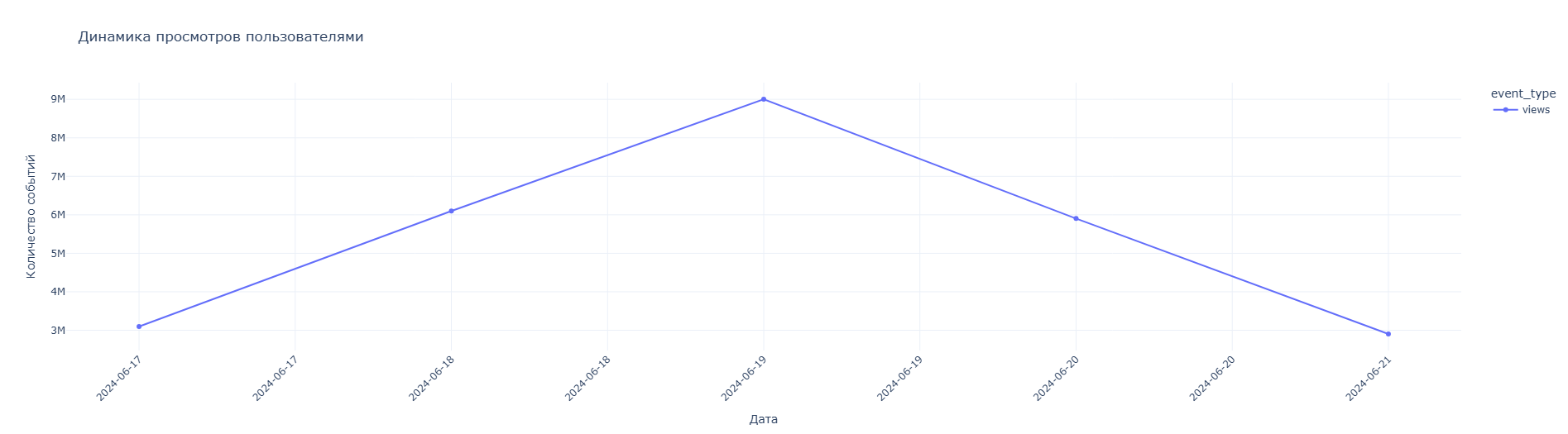
#### **Удержание клиентов (Retention rate)**

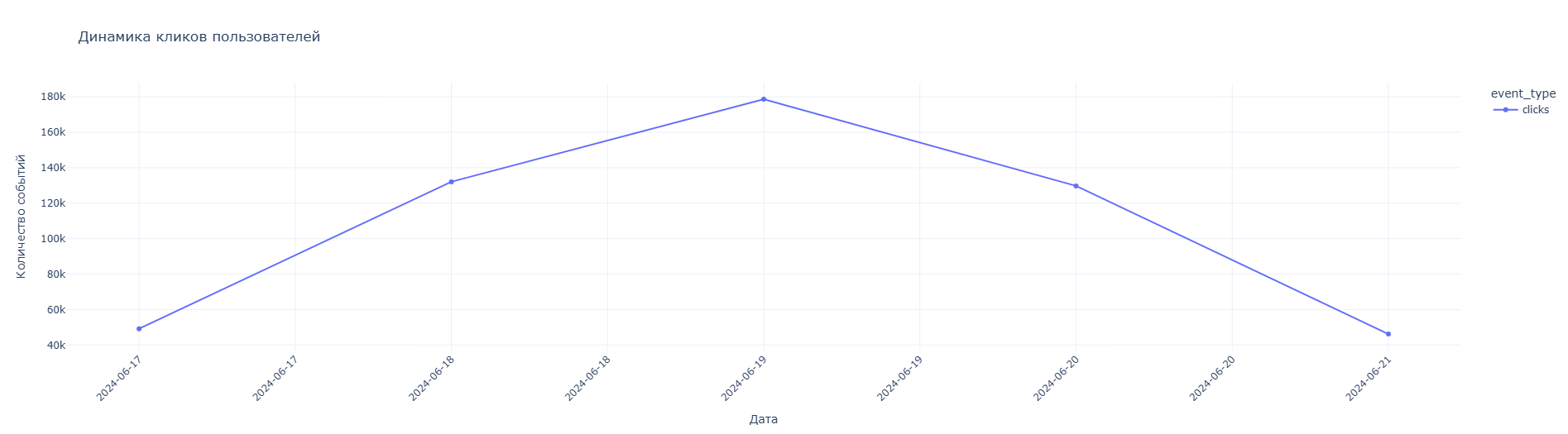


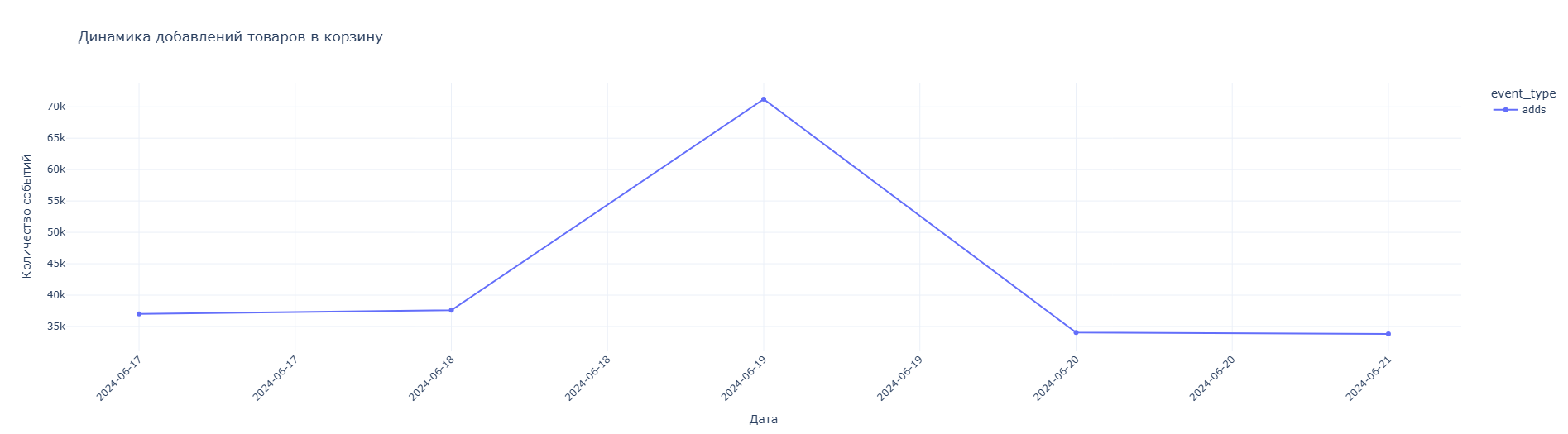
Когортный анализ удержания клиентов показывает, что до 70% теряется в течение первых 5 дней.

При этом каждая последующая когорта пользователей быстрее теряет интерес.

#### **Динамика активностей пользователей**







На 19.06.2024 приходится пик активности пользователей. В этот день совершено в разы больше добавлений товаров в корзину (70к) против 37к в остальные дни.

#### **User-based конверсия**

User-based конверсия показывает кол-во уникальных пользователей, совершивших хотя бы одно действие на каждом этапе.

Плюсы:

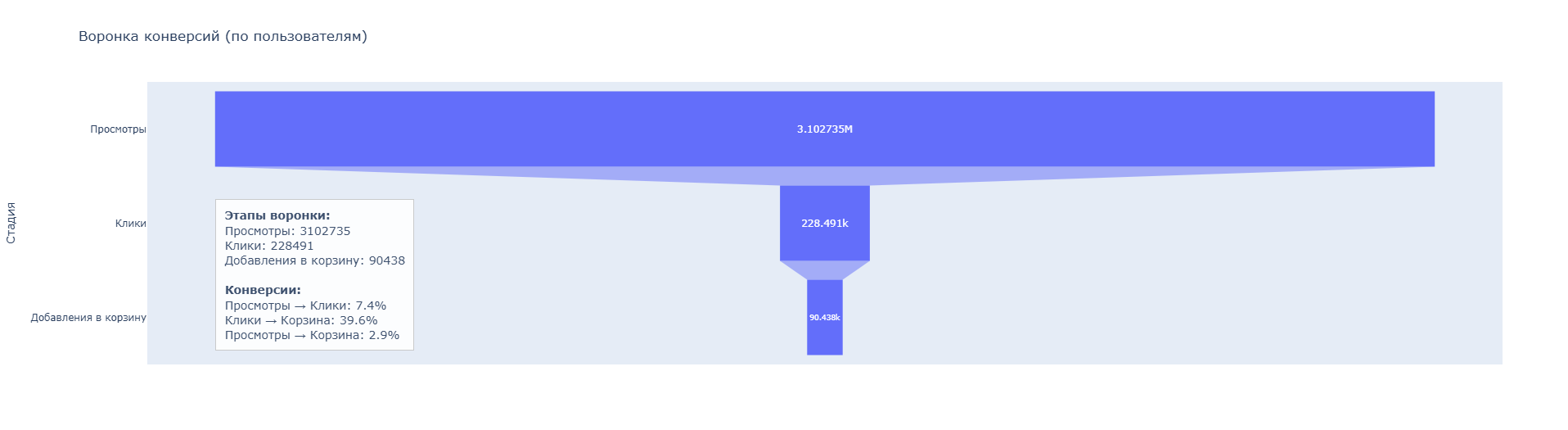
* Позволяет понять, сколько пользователей вовлечены на каждом этапе.
* Полезно, если фокус на поведении пользователей, а не на частоте действий.

Минусы:

* Один пользователь может кликнуть 10 раз, но он считается один раз.
* Не учитывается интенсивность активности.

Когда использовать:

Если хотим оценить воронку вовлечения пользователей, а не действия.



Конверсия в просмотр рекламы составляет 7,4%. Есть потенциал роста за счет подбора более таргетированных товаров, указанных в рекламе.

Целевая метрика - конверсия в добавление товара в корзину после просмотра рекламы равна 39,6%. После просмотра рекламы пользователи с хорошей конверсией добавляют товар в корзину.

#### **Event-based конверсия**

Event-based конверсия показывает все действия (просмотры, клики, добавления в корзину) на каждом этапе.

Плюсы:

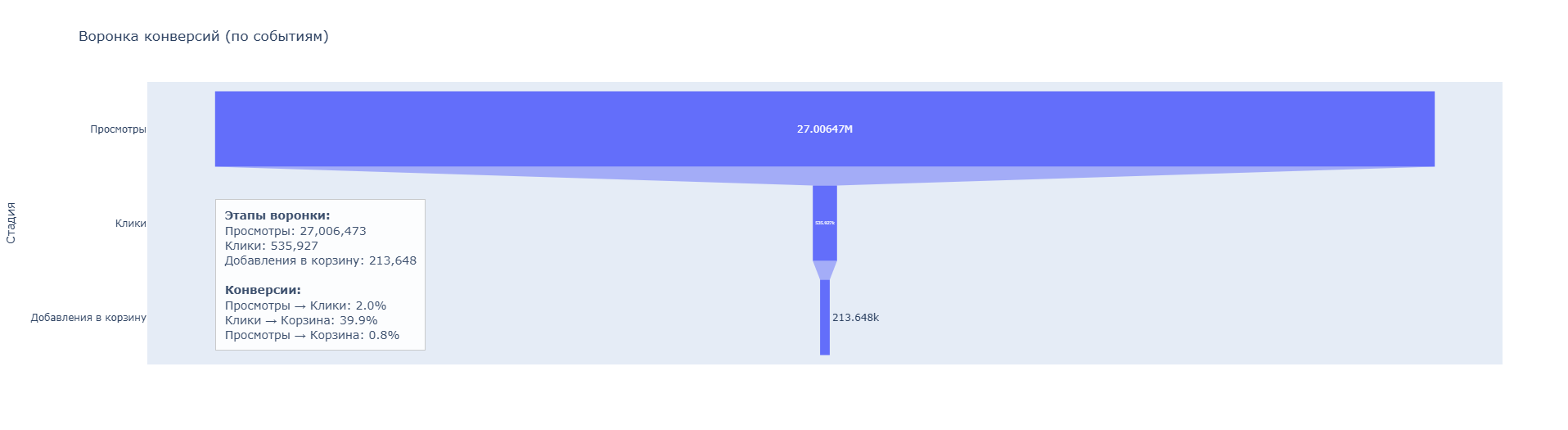
* Учитывает всю активность — даже если один пользователь кликнул 10 раз.
* Полезно для оценки эффективности контента, рекламы и интерфейса.

Минусы:

* Много шума — активные пользователи могут «зашумлять» статистику.

Когда использовать:

Когда важны взаимодействия с системой, кликабельность и производительность рекламы..

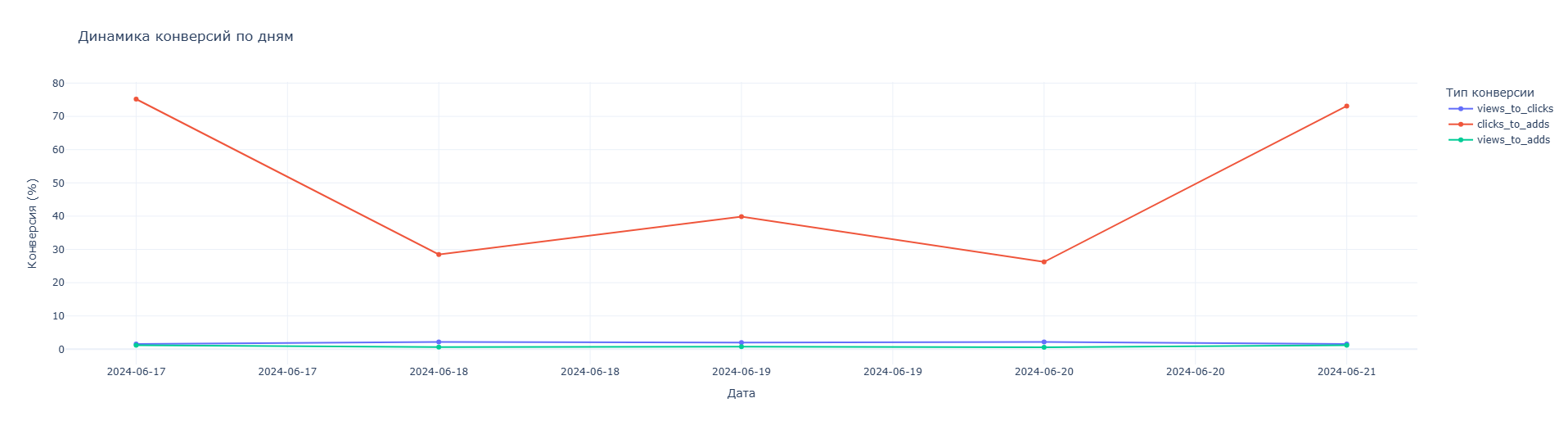


Конверсия в просмотр рекламы составляет 2,0%. Есть потенциал роста за счет подбора более таргетированных товаров, указанных в рекламе.

Целевая метрика - конверсия в добавление товара в корзину после просмотра рекламы равна 39,9%. После просмотра рекламы пользователи с хорошей конверсией добавляют товар в корзину.

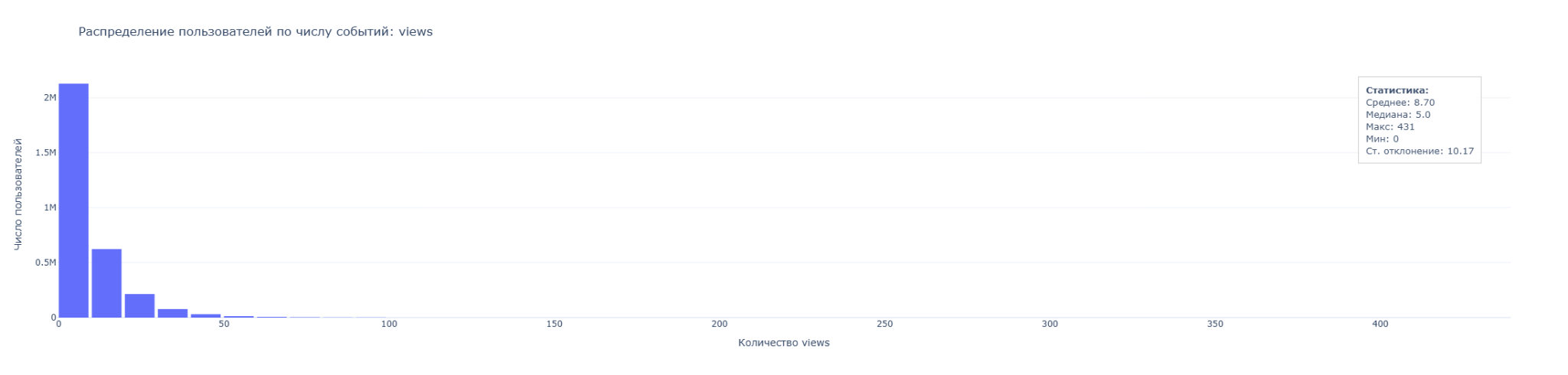
**Вывод**: Для расчета целевой метрики мы будем использовать **User-based конверсию,** так как нам важно оценить поведение пользователей.

#### **Динамика конверсий по дням**

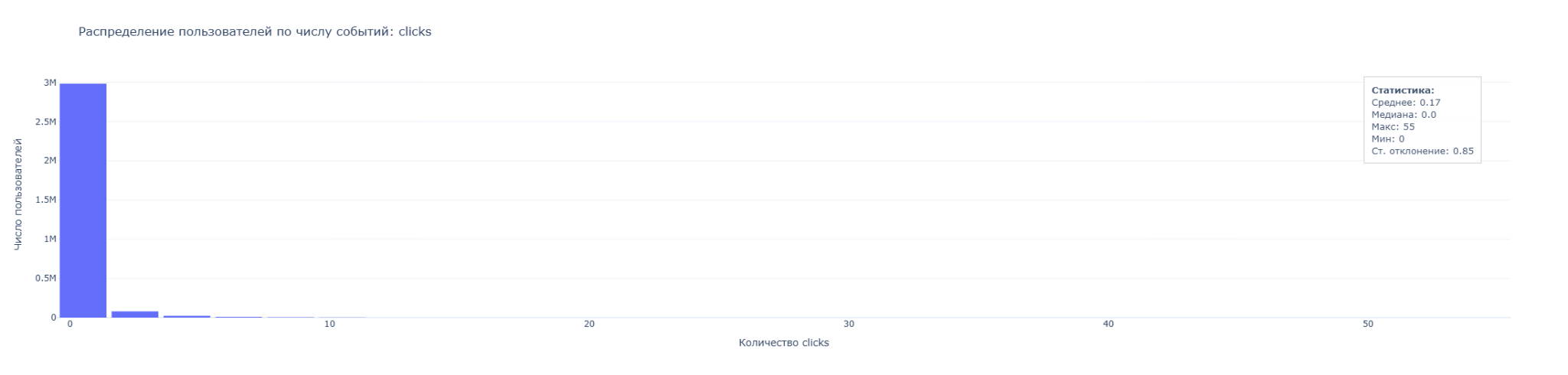


На 19.06.2024 целевая конверсия упала с 70% до 40%. В этот день наблюдалось большое кол-во пользователей при более низкой конверсии.

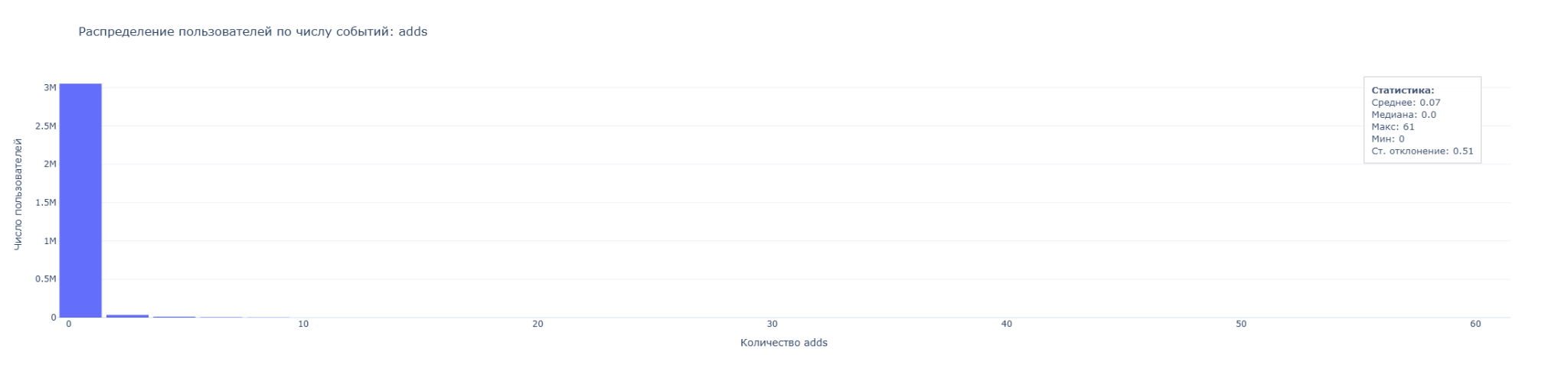
#### **Распределение пользователей по кол-ву активностей**



Большинство клиентов просмотрели главную страницу не более 30 раз.



Большинство клиентов просмотрели рекламу менее 3 раз.



Большинство клиентов добавили товары из рекламы в корзину менее 2 раз.

# **Определение MDE и минимальный размер аудитории**

#### **Расчет ключевой метрики для каждого уникального пользователя**

Для расчета целевой метрики необходимо создать сводную таблицу по дням и уникальным клиентам с булевыми полями**: было ли событие (view, click. add) у клиента в этот день.**

Далее вычисляется конверсия просмотра рекламы в добавление товара в корзину.

**Результат расчета (первые 5 строк):**

| **event\_type** | **client\_id** | **views** | **clicks** | **adds** | **CR** |
| --- | --- | --- | --- | --- | --- |
| **0** | 1 | 2 | 0 | 0 | NaN |
| **1** | 3 | 4 | 3 | 1 | 0.333333 |
| **2** | 4 | 4 | 0 | 0 | NaN |
| **3** | 19 | 4 | 0 | 0 | NaN |
| **4** | 27 | 2 | 0 | 0 | NaN |

#### **Расчет статистических показателей дисперсии и среднего значения ключевой метрики**

Для расчета MDE требуется определить дисперсию и среднее значение целевой метрики (CR).

Так как целевая метрика является отношением двух независимых величин, для определения дисперсии и среднего значения используется дельта-метод.

**Результат расчета:**  
Среднее значение CR: 0.4443

Дисперсия CR (дельта-метод): 0.334631

#### **Определение MDE (минимальный наблюдаемый эффект) и минимальной аудитории, необходимой для проведения теста.**

Для проведения теста необходимо определить размер аудитории, на которой будет проводиться тест.

Для проведения теста будет сформированы 2 гомогенные группы одинакового размера (тестовая и контрольная группы).

Минимальный размер аудитории зависит от MDE, который мы хотим заметить при проведении теста. Чем ниже MDE, тем большая аудитория должна быть задействована в тесте, чтобы желаемый эффект был обнаружен.

Мы не хотим привлекать большую аудиторию для тестирования, чтобы ограничить возможные негативные эффекты, в случае если гипотеза является неверной.

Поэтому максимальный размер аудитории мы устанавливаем в размере 20% от уникального количества пользователей.

**Результат расчета:**  
Чтобы добиться максимальной задействованной аудитории в 20%, нам нужно зафиксировать минимальный эффект, который A/B-тест сможет с высокой вероятностью обнаружить — это изменение конверсии (CR) на 0.0093 в абсолюте.

Это означает, что для заданных параметров теста мы можем уловить не менее чем ~2.08% относительное изменение метрики.

При этом Минимальный размер аудитории для одной группы при проведении А/Б теста составит 310 420.6 .

Минимальный размер аудитории для одной группы при проведении А/Б теста в % от кол-ва уникальных пользователей составит 10.00 %.

# **Анализ результатов А/Б теста**

#### **Структура данных результата теста**

Исходные данные содержат 981 510 строк и 11 столбцов.

При анализе данных было обнаружено и удалено 6 486 дубликатов.

Пропусков в исходных данных не обнаружено.

Количество уникальных клиентов в контрольной группе (группа А): 200 786.

Количество уникальных клиентов в тестовой группе (группа Б): 200 474.

Для корректного анализа результатов теста была проведена проверка на пересечение клиентов в контрольной и тестовой группах.

Количество клиентов, попавших в обе группы: 6 393. Клиенты, входящие в обе группы, были удалены.

**Результат расчета:**

Количество уникальных клиентов в контрольной группе: 194 393

Количество уникальных клиентов в тестовой группе: 194 081

Контрольная группа составляет: 62.6% от минимальной аудитории, необходимой для проведения теста.

Тестовая группа составляет: 62.5%, необходимой для проведения теста.

**Вывод: Проведенный тест не охватывает достаточной аудитории, чтобы обнаружить желаемый MDE.**

#### **Анализ метрик в контрольной и тестовой группах и определение статистической значимости полученных результатов (расчет конверсий ведется по клиентам и дням)**

При анализе итогов А/Б теста было выявлено, что рассчитать все метрики согласно дизайну теста не представляется возможным ввиду отсутствия необходимых данных: количество просмотров главной страницы. Поэтому из анализа были исключены метрики:

* Кликабельность секции рекламы (CTR) = количество кликов / количество просмотров главной страницы
* Конверсия в заказ от просмотра главной страницы

После расчета метрик был рассчитан [[4]](#footnote-3)uplift каждой метрики и определена статистическая значимость изменений по каждой метрике.

При определении статистической значимости для конверсионных метрик (CR клик → добавление, CR добавление → заказ, Процент выкупа корзины) применялся **z-test.**

При определении статистической значимости для метрик, основанных на средних значениях (Добавление в корзину/ Показ рекламы, Средний чек товаров в корзине, Заказ / Добавление в корзину, Средний чек, Заказов на клиент в день, Заказов на клиента, Выручка на клиента(ARPU)) применялся **t-test**.

**Результат расчета:**

| **Z/T-статистика** | **P-value** | **Стат. значимость** | **Тип теста** |
| --- | --- | --- | --- |
| -0.9868 | 0.3238 | ⚠️ | z-test |
| 1.9788 | 0.0478 | **✅** | z-test |
| 48.4790 | 0.0000 | ✅ | z-test |
| -0.9034 | 0.3663 | ⚠️ | t-test |
| 0.0641 | 0.9489 | ⚠️ | t-test |
| 0.1151 | 0.9084 | ⚠️ | t-test |
| -1.2022 | 0.2293 | ⚠️ | t-test |
| -2.1650 | 0.0304 | ✅ | t-test |
| -1.9956 | 0.0460 | ✅ | t-test |
| -1.9956 | 0.0460 | ✅ | t-test |

**Вывод:**

1. Ключевая метрика CR клик → добавление за время проведения не прокрасилась. Разница метрики в контрольной и тестовой группах статистически не значима.
2. А/Б тест не выявил значимой статистической разницы по ключевой метрике и не может считаться успешным. Так как изначально была выявлена недостаточность в выборке теста, предлагается продолжить тестирование, чтобы набрать минимальную аудиторию для выявления MDE.
3. При этом изменения в вспомогательных метриках (CR добавление → заказ, Процент выкупа корзины, Заказов на клиента в день, Заказов на клиента, Выручка на клиента) оказались статистически значимы.

#### **Анализ изменения поведения пользователей при проведении теста**

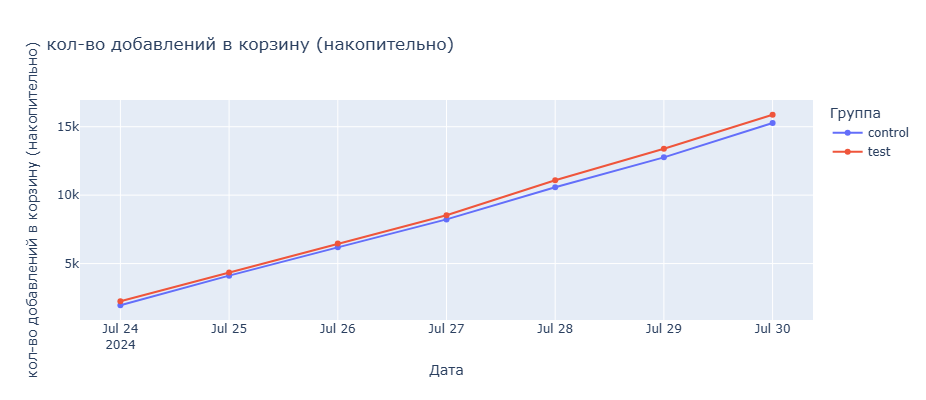
**Анализ метрик**

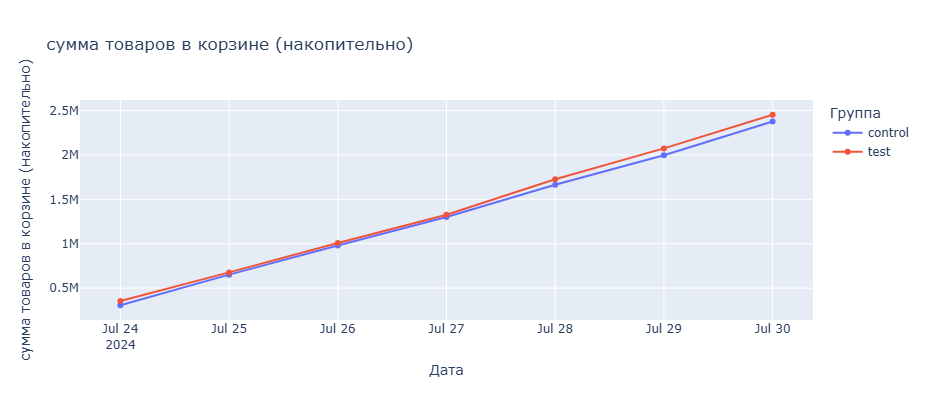
| **Метрика** | **Control** | **Test** | **Uplift, %** | **Z/T-статистика** | **P-value** | **Стат. значимость** | **Тип теста** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| CR клик → добавление | 0.2813 | 0.2773 | -1.4208 | -0.9868 | 0.3238 | ⚠️ | z-test |
| CR добавление → заказ | 0.8754 | 0.8863 | 1.2509 | 1.9788 | 0.0478 | **✅** | z-test |
| Процент выкупа корзины | 0.6588 | 0.6795 | 3.1512 | 48.4790 | 0.0000 | ✅ | z-test |
| Добавление в корзину/ Показ рекламы | 0.3675 | 0.3785 | 3.0098 | -0.9034 | 0.3663 | ⚠️ | t-test |
| Средний чек товаров в корзине | 155.565 | 154.474 | -0.7013 | 0.0641 | 0.9489 | ⚠️ | t-test |
| Заказ / Добавление в корзину | 0.3597 | 0.3658 | 1.7148 | 0.1151 | 0.9084 | ⚠️ | t-test |
| Средний чек | 284.941 | 286.939 | 0.7009 | -1.2022 | 0.2293 | ⚠️ | t-test |
| Заказов на клиента в день | 0.0117 | 0.0123 | 5.3655 | -2.1650 | 0.0304 | ✅ | t-test |
| Заказов на клиента | 0.0283 | 0.0299 | 5.8660 | -1.9956 | 0.0460 | ✅ | t-test |
| Выручка на клиента | 8.0546 | 8.5868 | 6.6081 | -1.9956 | 0.0460 | ✅ | t-test |

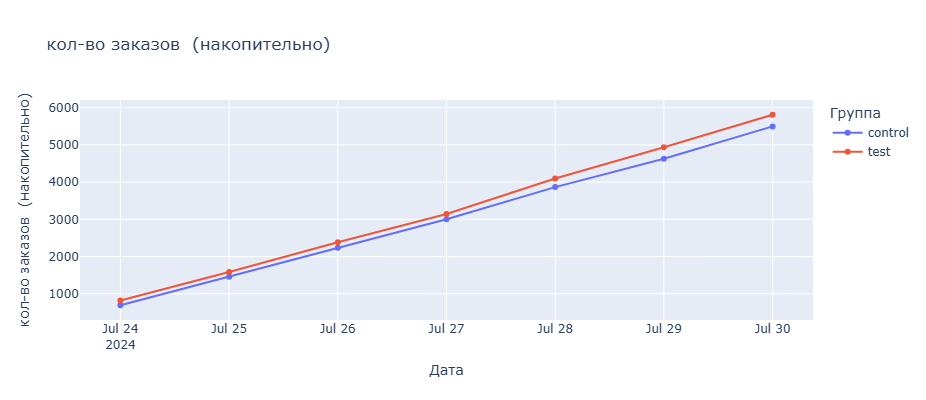
**Вывод:**

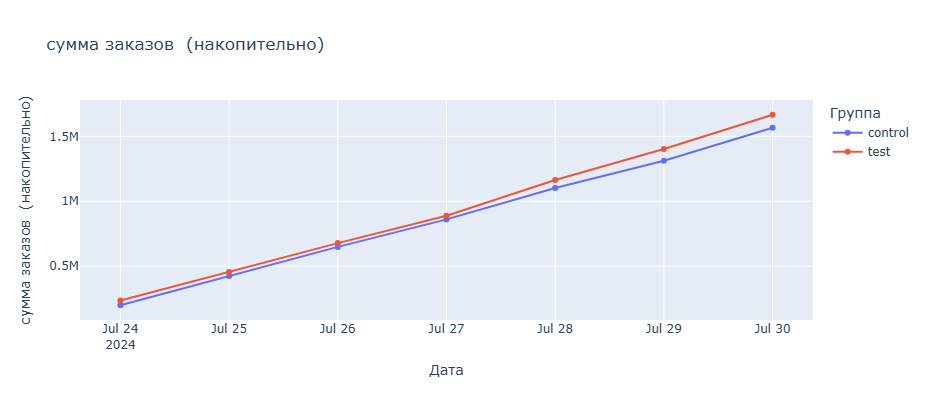
1. В тестовой и контрольной группах снизилась конверсия в добавление товаров в корзину после просмотра рекламы по сравнению с исходными данными с 39,9% до 27,7% и 28,1% соответственно. При этом конверсия в добавление товаров в корзину после просмотра рекламы при event-based подходе осталась примерно на том же уровне - 37,9% и 36,8% соответственно. Можно сделать вывод о том, что доля пользователей, готовых купить товар после просмотра рекламы снизилась, а активность таких людей (кол-во кликов и добавление товара в корзину) выросла.
2. Конверсия в покупку после добавления товара в корзину составляет 88,6% и 87,5% для тестовой и контрольной групп соответственно. Конверсия в покупку после добавления товаров в корзину при event-based подходе равна 36,6% и 36,0% для тестовой и контрольной групп соответственно. Можно сделать вывод о том, что большинство пользователей делает одну покупку после добавления товара в корзину. Однако, вероятность последующих покупок после добавления товара в корзину резко падает. При этом процент выкупа корзины (по стоимости товаров) составляет - 67,9% и 65,9%, что говорит о том, что пользователи чаще отказываются от покупки товаров, добавленных в корзину, с более низким средним чеком.
3. Средний чек товаров в корзине составляет 154,5 и 155,6 у.е. для тестовой и контрольной групп соответственно. Средний чек в заказе составляет 286,9 и 284,9 у.е. соответственно. Средний чек в заказе выше среднего чека корзины, что подтверждает выводы из п. 2.
4. Среднее количество заказов на клиента - 0,0299 и 0,0283 соответственно
5. Выручка на клиента (ARPU) - 8,59 и 8,05 у.е.

**Динамика изменения исходных показателей и разница в тестовой и контрольных группах. (накопительно)**



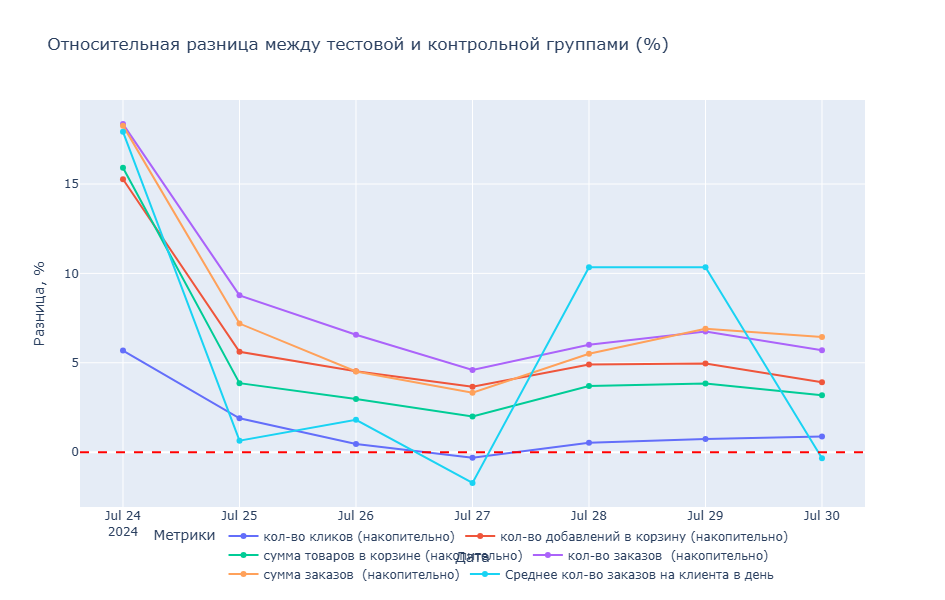






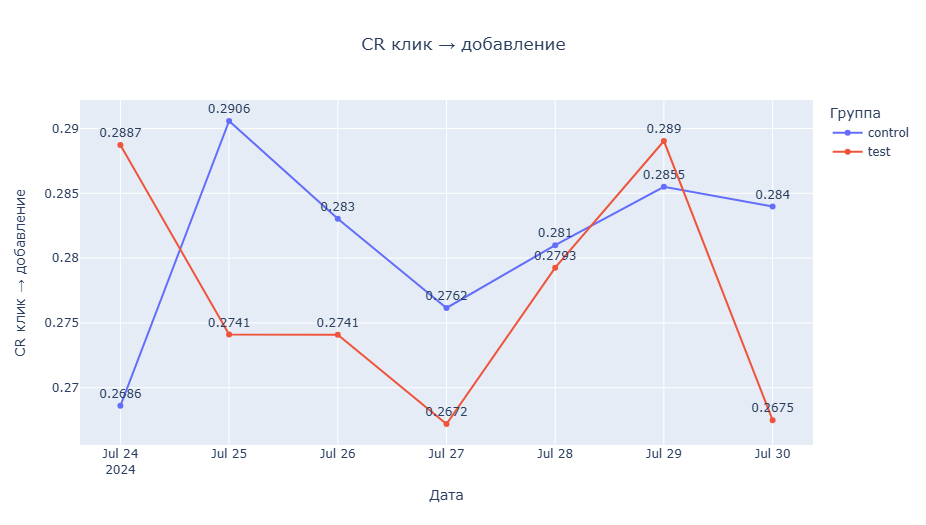


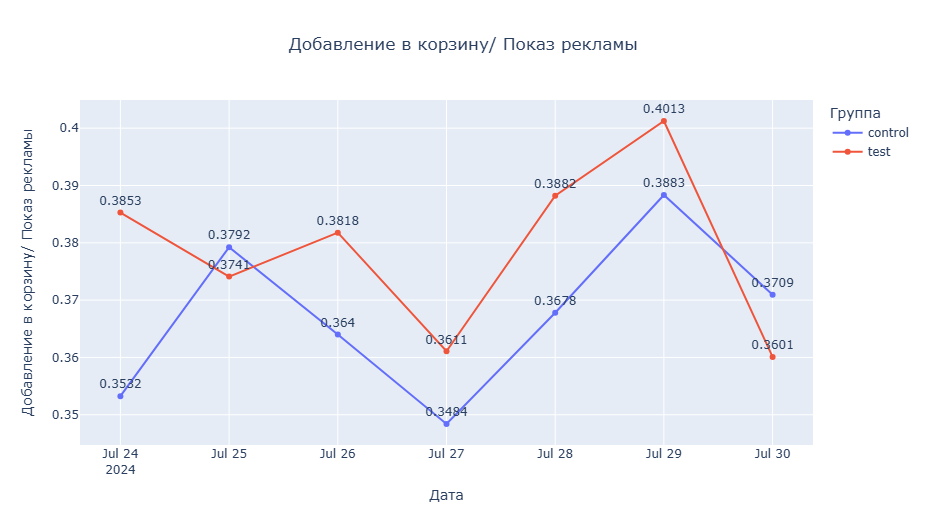
Относительная разница между тестовой и контрольной группами (%)



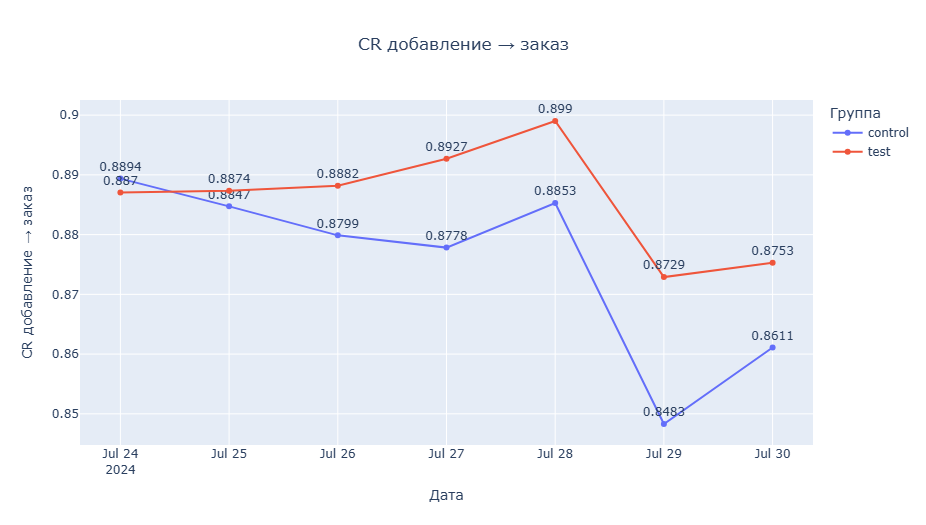
**Вывод:** Практически по всем показателям (клики, добавление в корзину, кол-во заказов, сумма заказов, сумма товаров в корзине) тестовая группа обогнала контрольную.

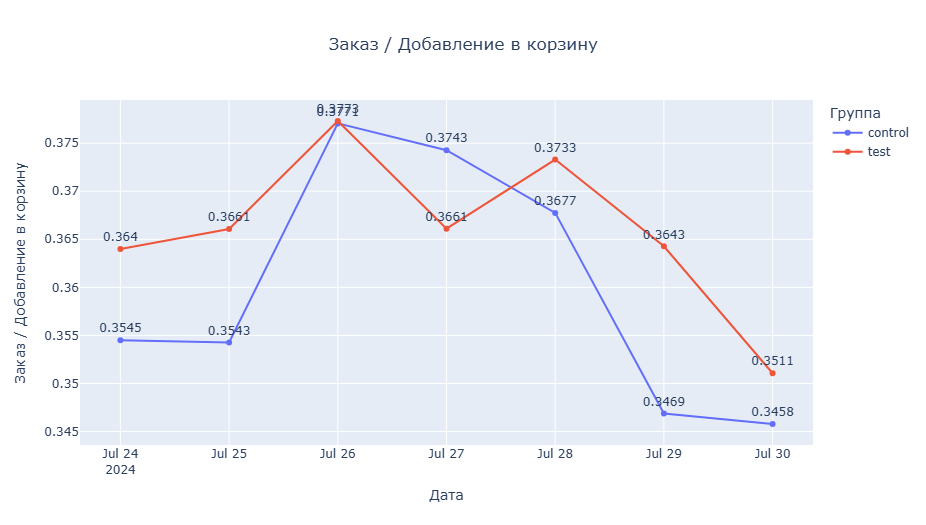
**Динамика изменения ключевых и вспомогательных показателей и разница в тестовой и контрольных группах.**

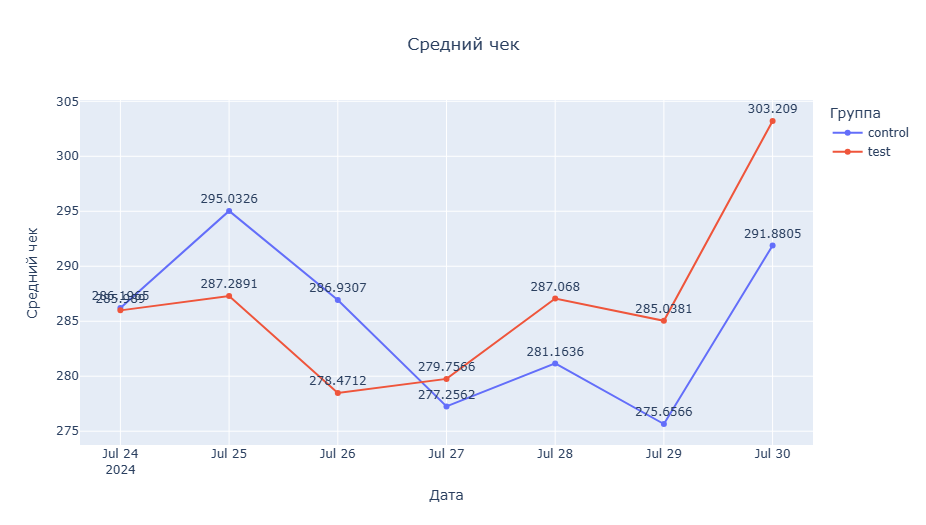


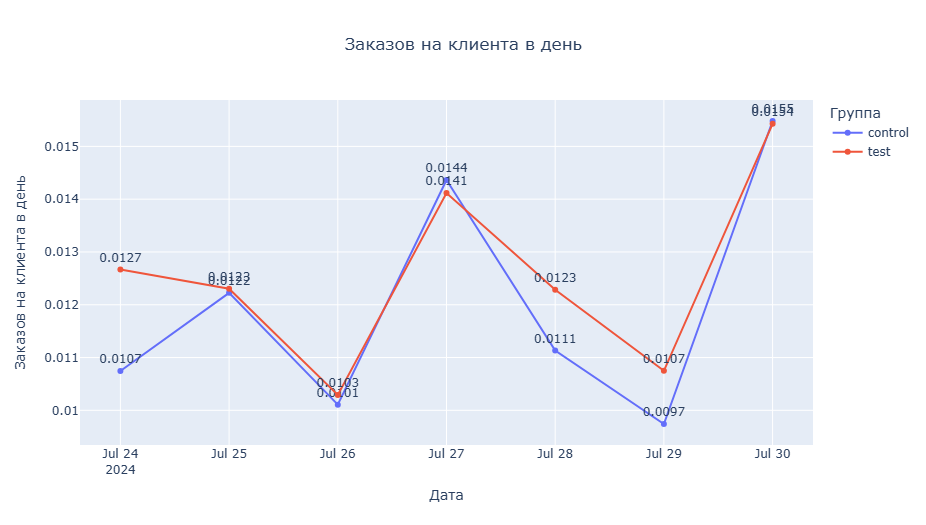


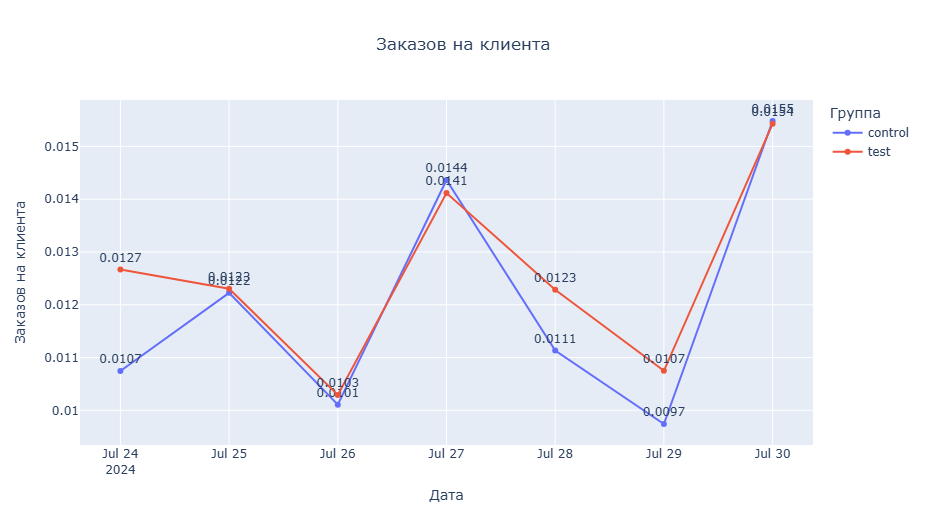


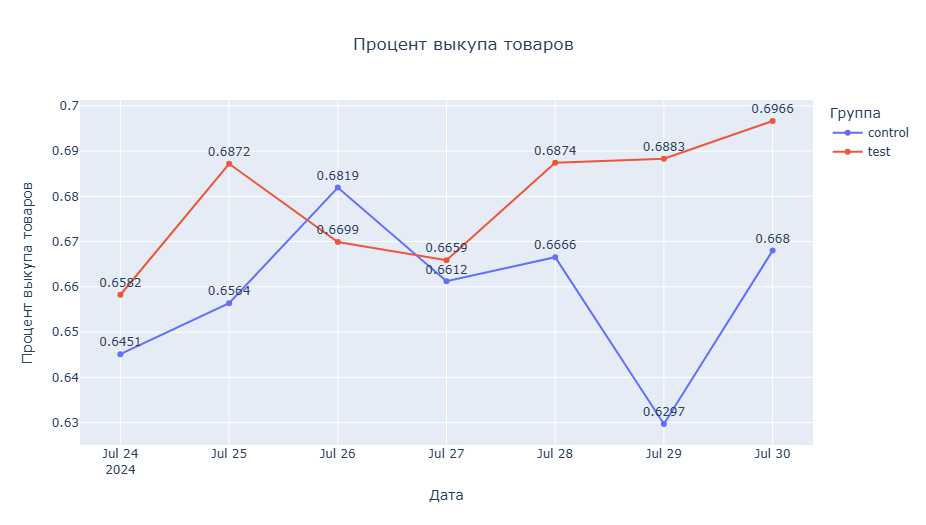






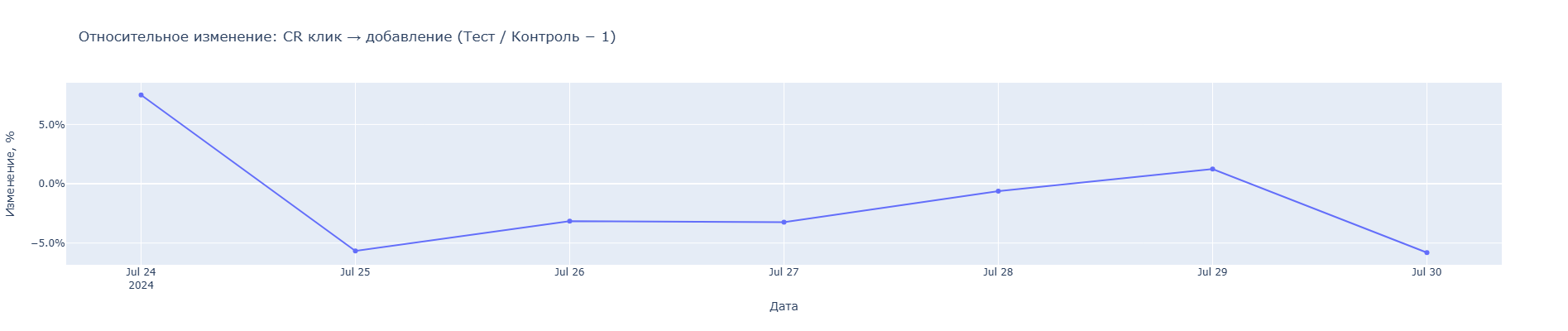




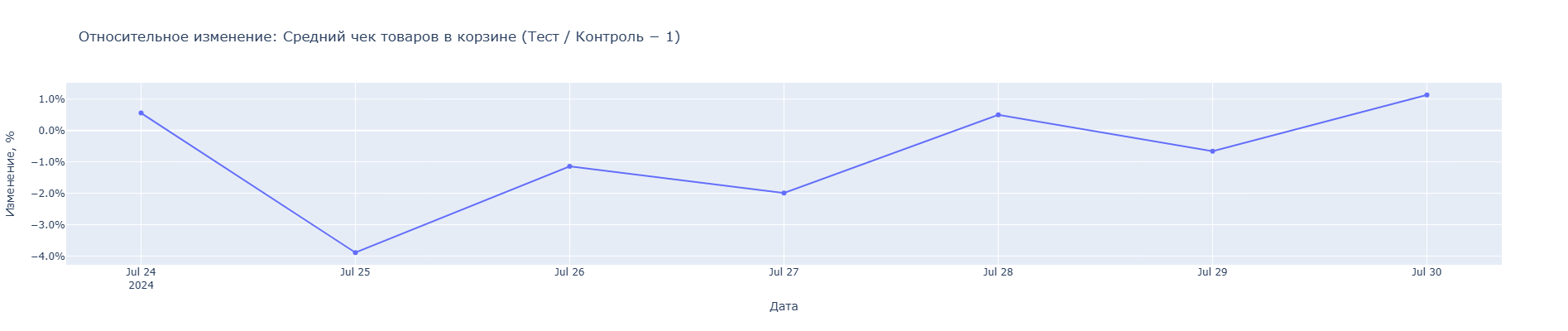


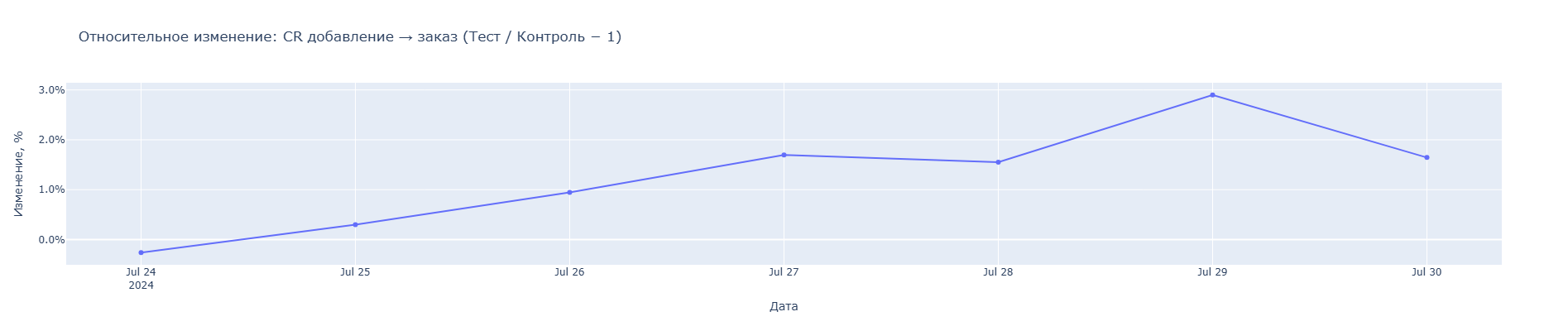


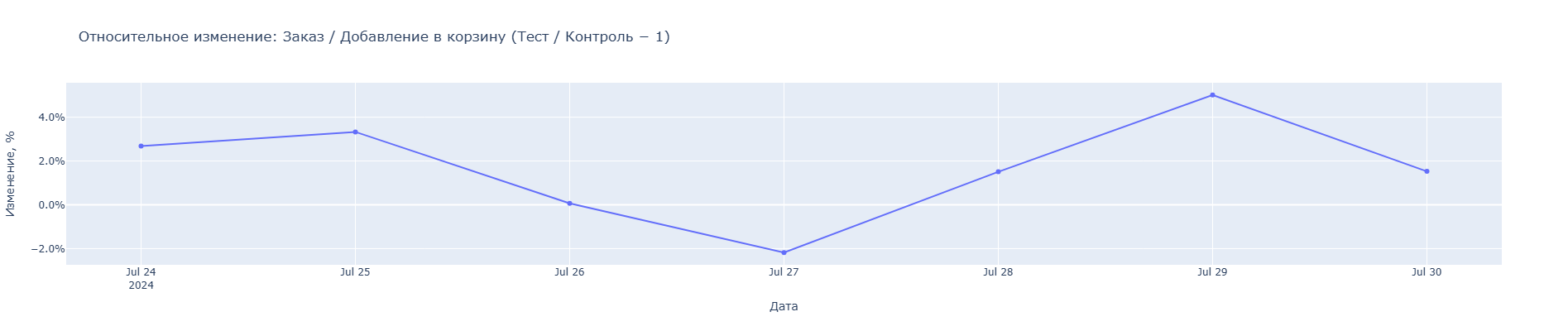
Относительная разница между тестовой и контрольной группами (%)

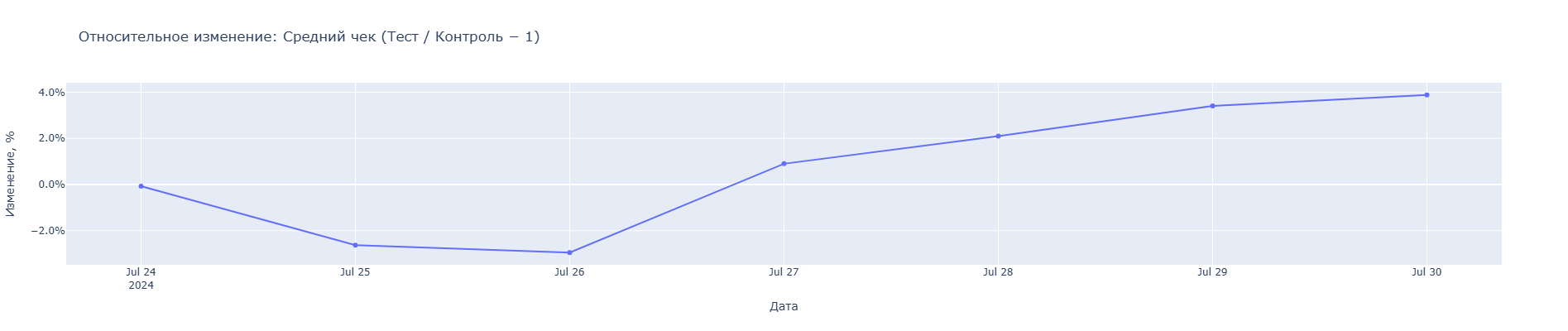


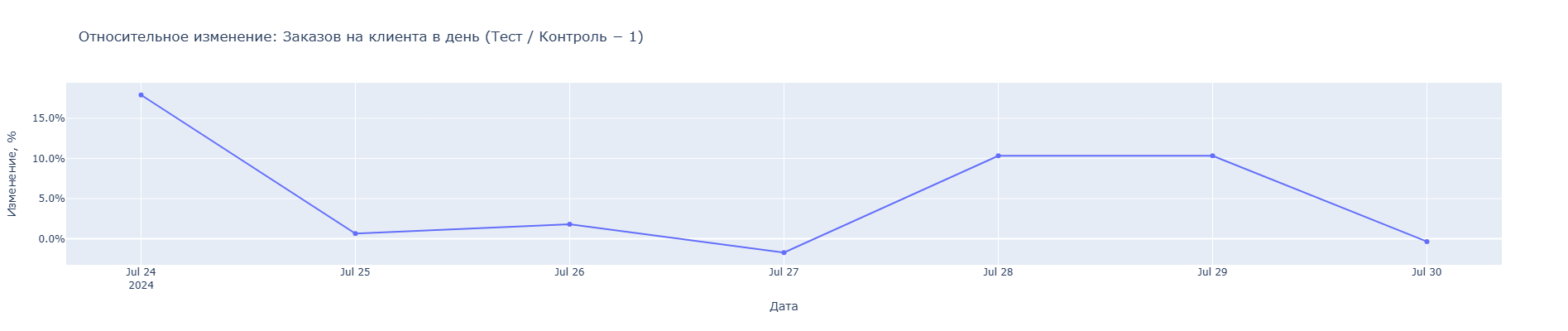


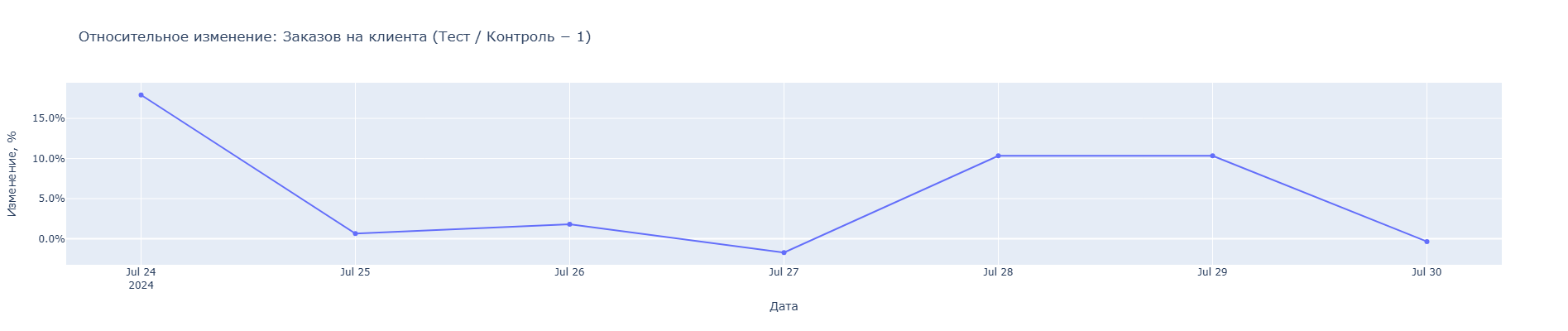


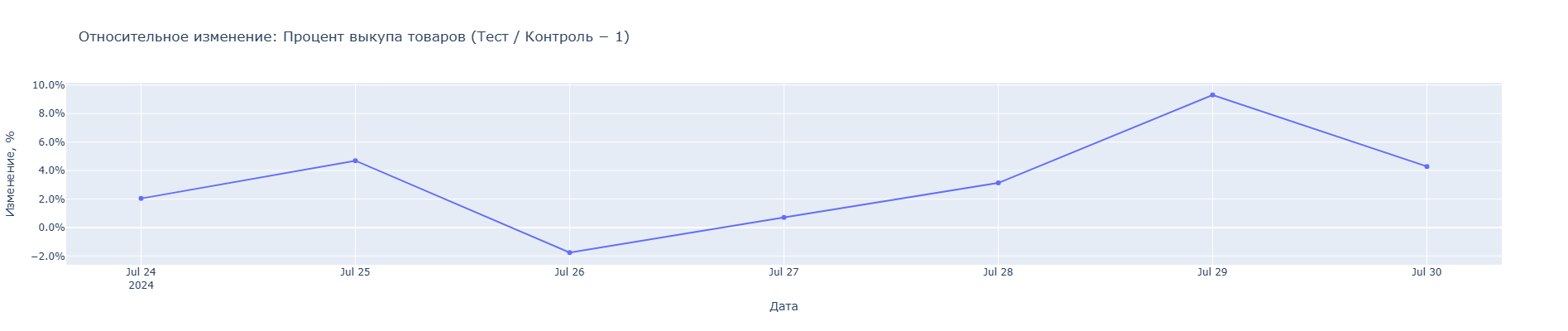


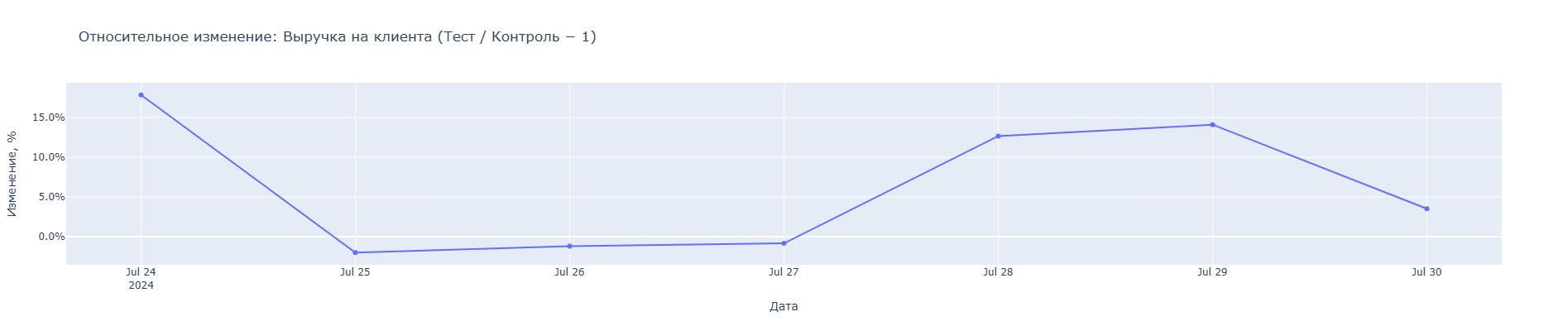












# **Заключение**

Перед нами стояла задача проверка гипотезы, что изменение размещения рекламы на странице маркетплейса влияет на взаимодействие пользователей с рекламой. Проверка гипотезы осуществлялась через проведение А/Б теста. Расчеты и визуализация проводились в notebook Google Collab (на языке Python). При этом были использованы библиотеки: pandas, numpy, plotly.express, seaborn, matplotlib.pyplot, scipy.stats.

Анализ результатов А/Б теста показал, что тест не выявил статистической значимости в разнице показателей контрольной и тестовой группы для целевой метрики CR клик → добавление. Кроме того, выборка кол-ва уникальных клиентов в тестовой и контрольных группах оказалась недостаточной (количество уникальных клиентов в контрольной группе (группа А): 200 786, количество уникальных клиентов в тестовой группе (группа Б): 200 474 при расчетной минимальной аудитории в 310 42,.6).

**Требуется продолжение проведения тестирования, так как проведенный тест не охватывает достаточной аудитории, чтобы обнаружить желаемый MDE.**

Из предложенных вспомогательных метрик, CR добавление → заказ, Процент выкупа корзины, Заказов на клиента в день, Заказов на клиента, Выручка на клиента оказались статистически значимы.

Из анализа поведения пользователей сделаны следующие выводы:

1. В тестовой и контрольной группах снизилась конверсия в добавление товаров в корзину после просмотра рекламы по сравнению с исходными данными с 39,9% до 27,7% и 28,1% соответственно. При этом конверсия в добавление товаров в корзину после просмотра рекламы при event-based подходе осталась примерно на том же уровне - 37,9% и 36,8% соответственно. Можно сделать вывод о том, что доля пользователей, готовых купить товар после просмотра рекламы снизилась, а активность таких людей (кол-во кликов и добавление товара в корзину) выросла.
2. Конверсия в покупку после добавления товара в корзину составляет 88,6% и 87,5% для тестовой и контрольной групп соответственно. Конверсия в покупку после добавления товаров в корзину при event-based подходе равна 36,6% и 36,0% для тестовой и контрольной групп соответственно. Можно сделать вывод о том, что большинство пользователей делает одну покупку после добавления товара в корзину. Однако, вероятность последующих покупок после добавления товара в корзину резко падает. При этом процент выкупа корзины (по стоимости товаров) составляет - 67,9% и 65,9%, что говорит о том, что пользователи чаще отказываются от покупки товаров, добавленных в корзину, с более низким средним чеком.
3. Средний чек товаров в корзине составляет 154,5 и 155,6 у.е. для тестовой и контрольной групп соответственно. Средний чек в заказе составляет 286,9 и 284,9 у.е. соответственно. Средний чек в заказе выше среднего чека корзины, что подтверждает выводы из п. 2.
4. Среднее количество заказов на клиента - 0,0299 и 0,0283 соответственно
5. Выручка на клиента (ARPU) - 8,59 и 8,05 у.е.

1. Гипотеза двусторонняя, так как мы не знаем, как изменение интерфейса повлияет на поведение пользователей. [↑](#footnote-ref-0)
2. Расчет ведется по флагу совершения хотя бы 1 действия уникальным клиентом в день [↑](#footnote-ref-1)
3. Расчет ведется по флагу совершения хотя бы 1 действия уникальным клиентом в день [↑](#footnote-ref-2)
4. В контексте A/B-тестирования (сплит-тестирования) uplift (прирост) — это метрика, которая показывает, насколько изменение (вариант B) улучшило ключевой показатель (например, конверсию, клики, продажи) по сравнению с исходной версией (вариант A). Это количественная оценка эффекта от внедрения изменений. [↑](#footnote-ref-3)