

Выделение групп пользователей на
основе поведения

Мобильные приложение —

 **НУЖНЫЕ
ВЕЩИ**

Цель – Анализ действий пользователей.

Оптимизация приложения и повышение в целевое событие.

Задача:

Выделите группы пользователей, которые различаются по метрикам:

- 1.retention rate,
- 2.время, проведённое в приложении,
- 3.частота действий,
- 4.конверсия в целевое действие — просмотр контактов.

- Проведите исследовательский анализ данных
- Сегментируйте пользователей на основе действий
- Проверьте статистические гипотезы

*1.Некоторые пользователи установили приложение по ссылке из yandex, другие — из google.
Проверьте гипотезу: две эти группы демонстрируют разную конверсию в просмотры контактов.*

*2.Сформулируйте собственную гипотезу. Дополните её нулевой и альтернативной гипотезами.
Проведите статистический тест.*

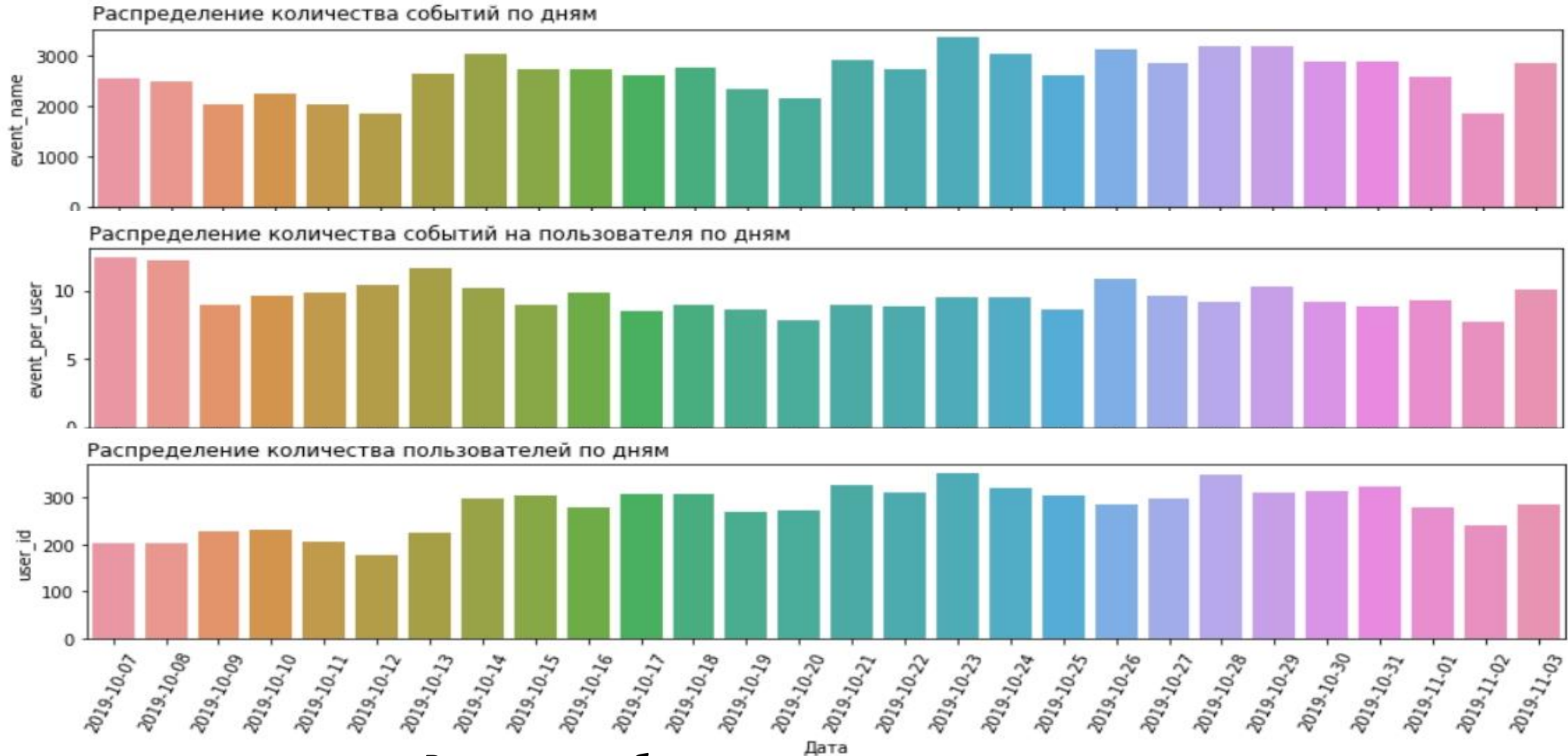
Данные

- Начало периода: 2019-10-07 00:00:00
- Окончание периода: 2019-11-03 23:58:12
- Продолжительность периода: 27 days 23:58:12
- Среднее количество сессий на пользователя: 2.42
- Среднее количество действий в рамках 1 сессии: 7.16
- Кол-во пользователей совершивших целевое действие: 981
- Что составляет: 22.85 % от общего кол-ва пользователей

СОДЕРЖАНИЕ

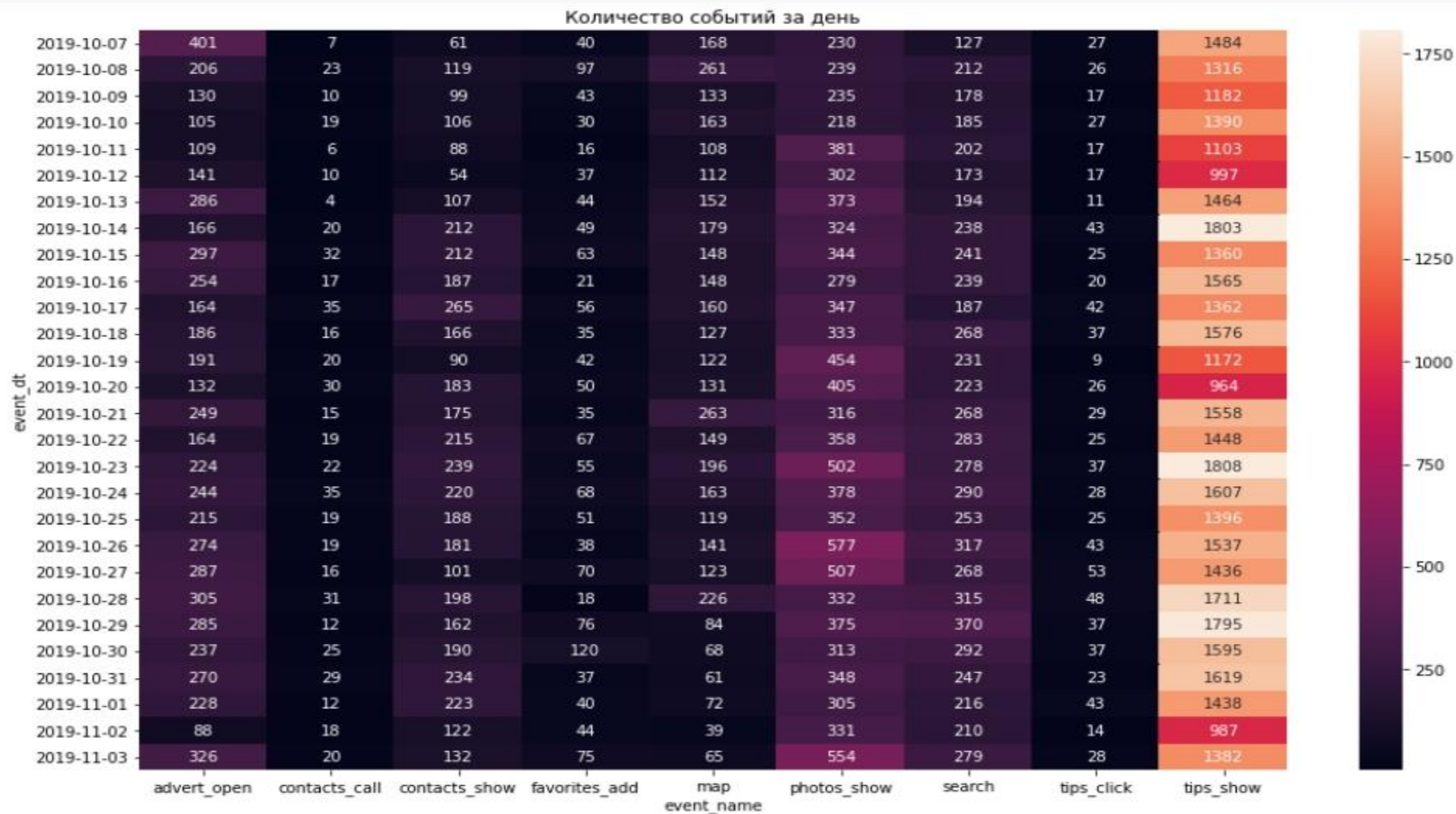
1. <u>Распределение количества событий и пользователей по времени</u>	<u>5</u>
2. <u>Тепловая карта</u>	<u>6</u>
3. <u>Гистограмма по дате и времени</u>	<u>7</u>
4. <u>Распределение событий за одну сессию</u>	<u>8</u>
5. <u>Частота событий</u>	<u>9</u>
6. <u>События по числу пользователей</u>	<u>10</u>
7. <u>Воронки событий</u>	<u>11</u>
8. <u>Сегментация пользователей</u>	<u>12</u>
9. <u>Ответы на вопросы заказчика</u>	<u>13</u>
10. <u>Retention Rate</u>	<u>14</u>
11. <u>Различие времени между распространенными событиями пользователей</u>	<u>15</u>
12. Статистические гипотезы: <u>Установка приложений из yandex и из google</u>	<u>16</u>
13. <u>При увеличении действий разных групп пользователей RR различается</u>	<u>17</u>
14. <u>РЕКОМЕНДАЦИИ</u>	<u>18</u>

Распределение количества событий и пользователей по времени

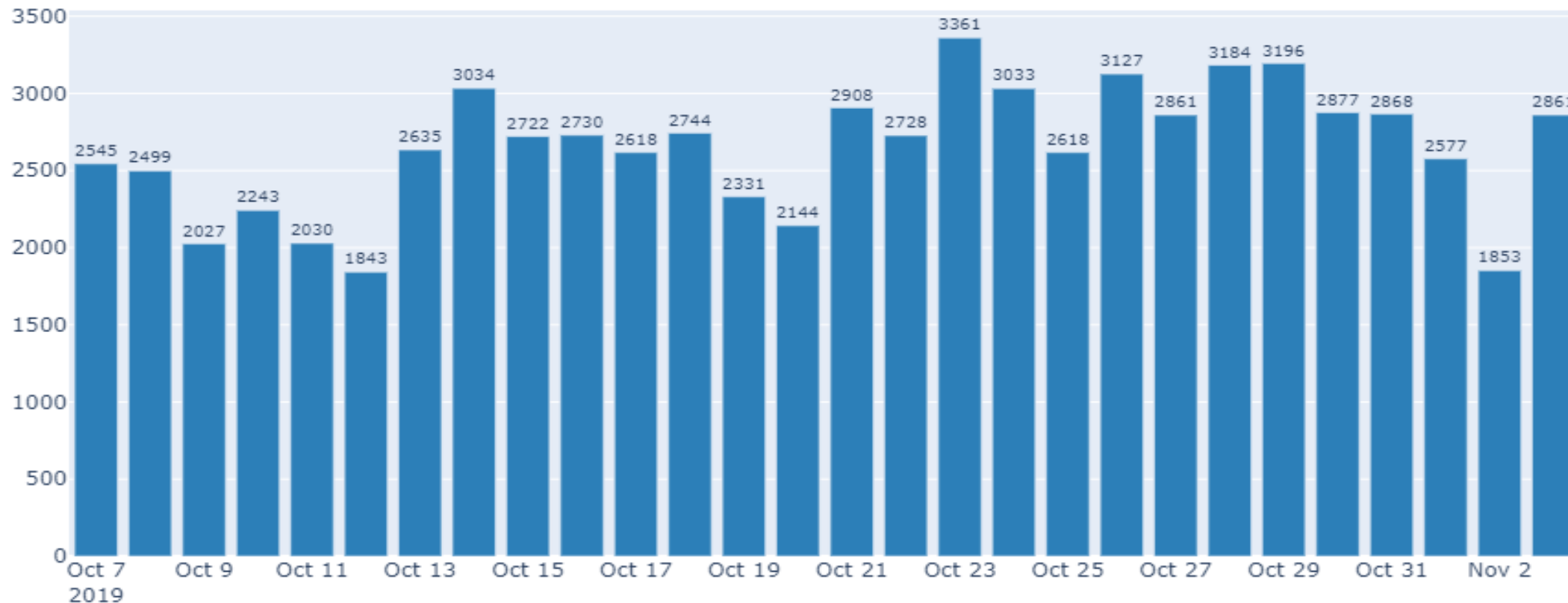


Резкие колебания в значениях отсутствуют

Тепловая карта

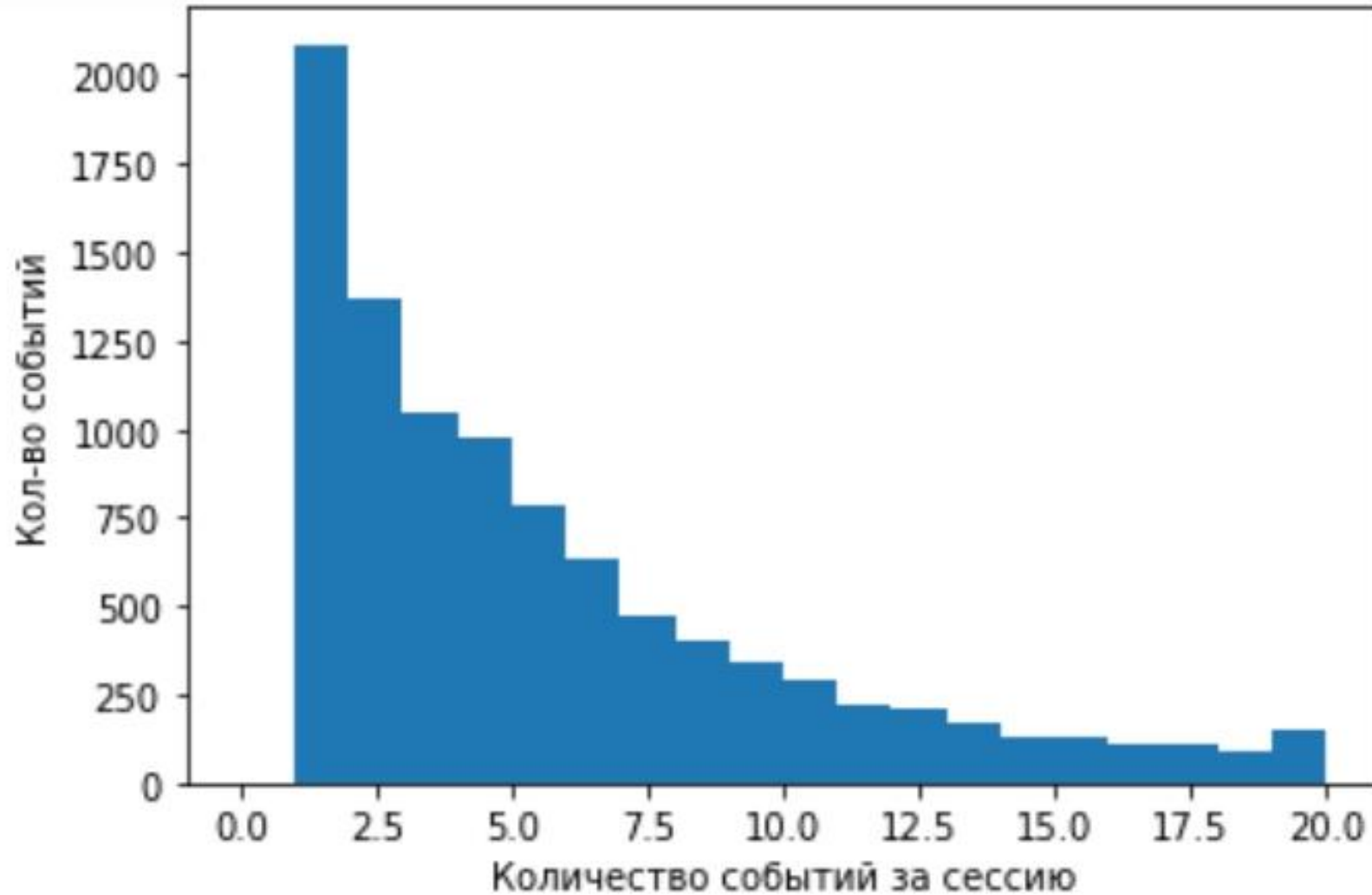


Гистограмма по дате и времени



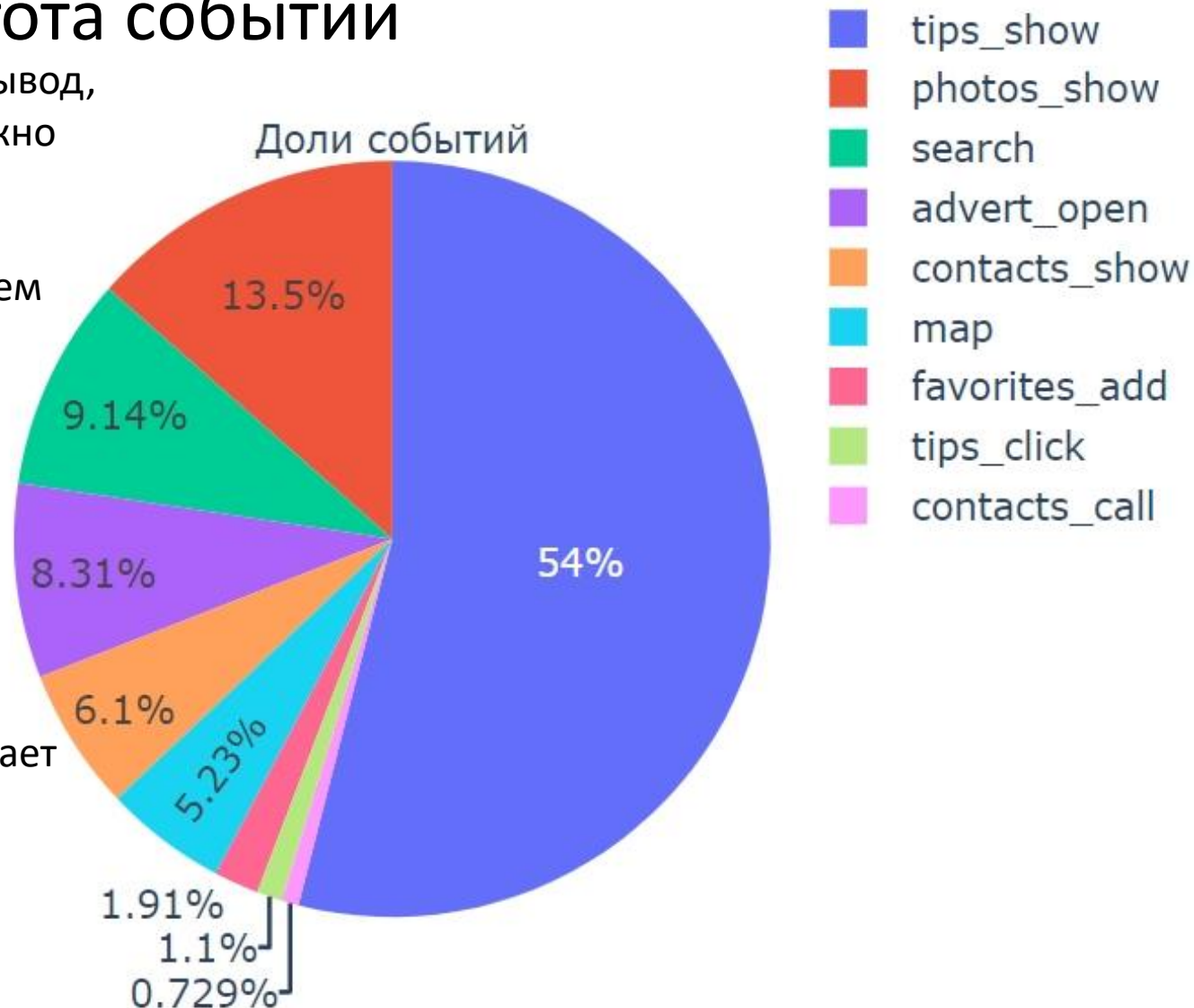
Данные достаточно целостные, без сильных провалов.

Распределение событий за одну сессию

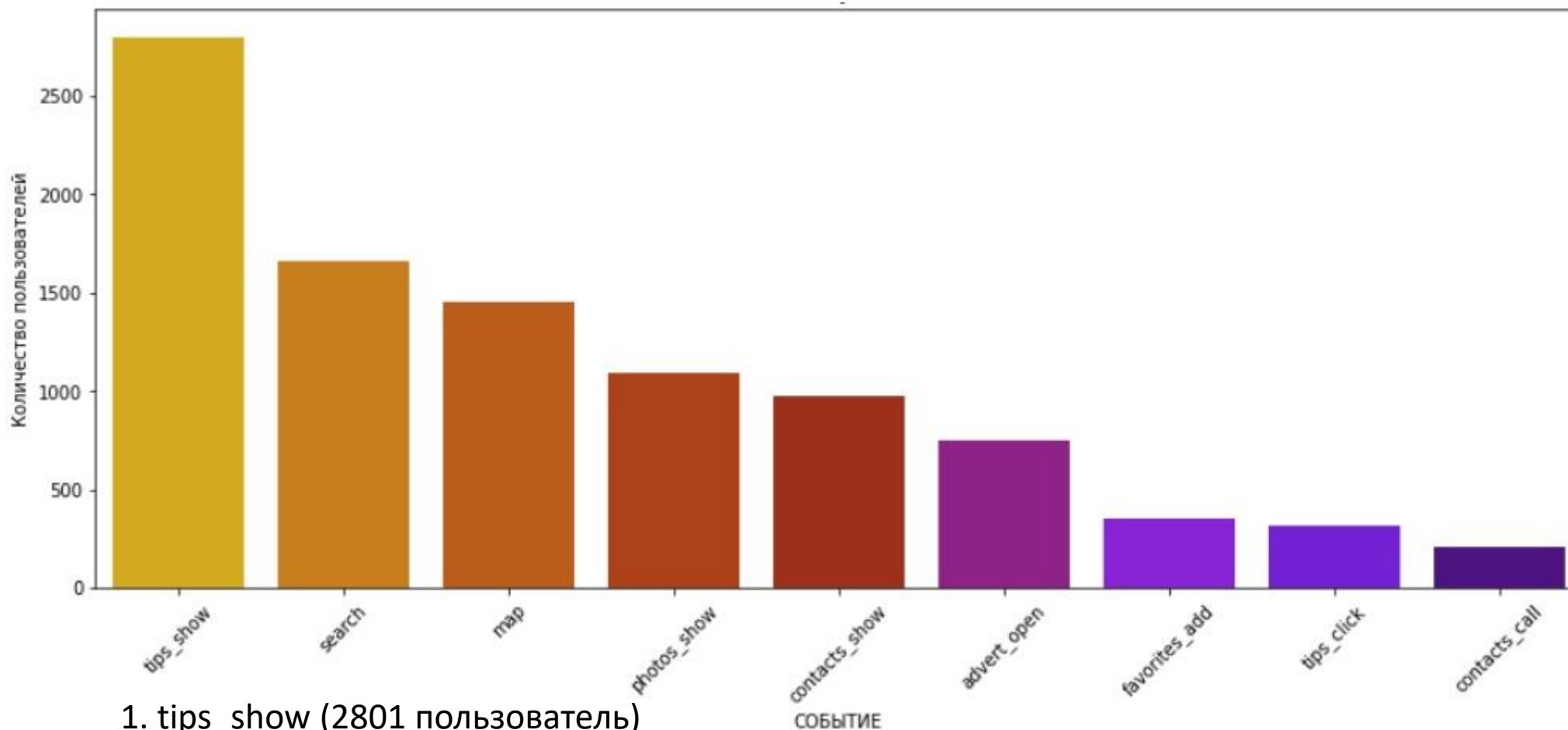


Частота событий

- С высокой долей вероятности можно сделать вывод, что начальным событием большинства сессий можно считать tips_show, так как доля этого события составляет 54 % от общего числа всех событий.
- Важно отметить, что tips_show показывается всем пользователям автоматически и не зависит от его действий. Его можно не учитывать в некоторых расчетах данного исследования.
- К сожалению, в приложении пока нет события "Появление главного экрана", которое помогло бы нам четко дифференцировать начальный этап сессии каждого пользователя, этот момент можно учесть в качестве рекомендации разработчикам.
- Среди оставшихся событий второе место занимает photos_show (13,5%), 3 и 4 места, соответственно, search (9,14%) и наше ЦС advert_open (8,31%).
- Прочие события занимают меньшие доли.



События по числу пользователей

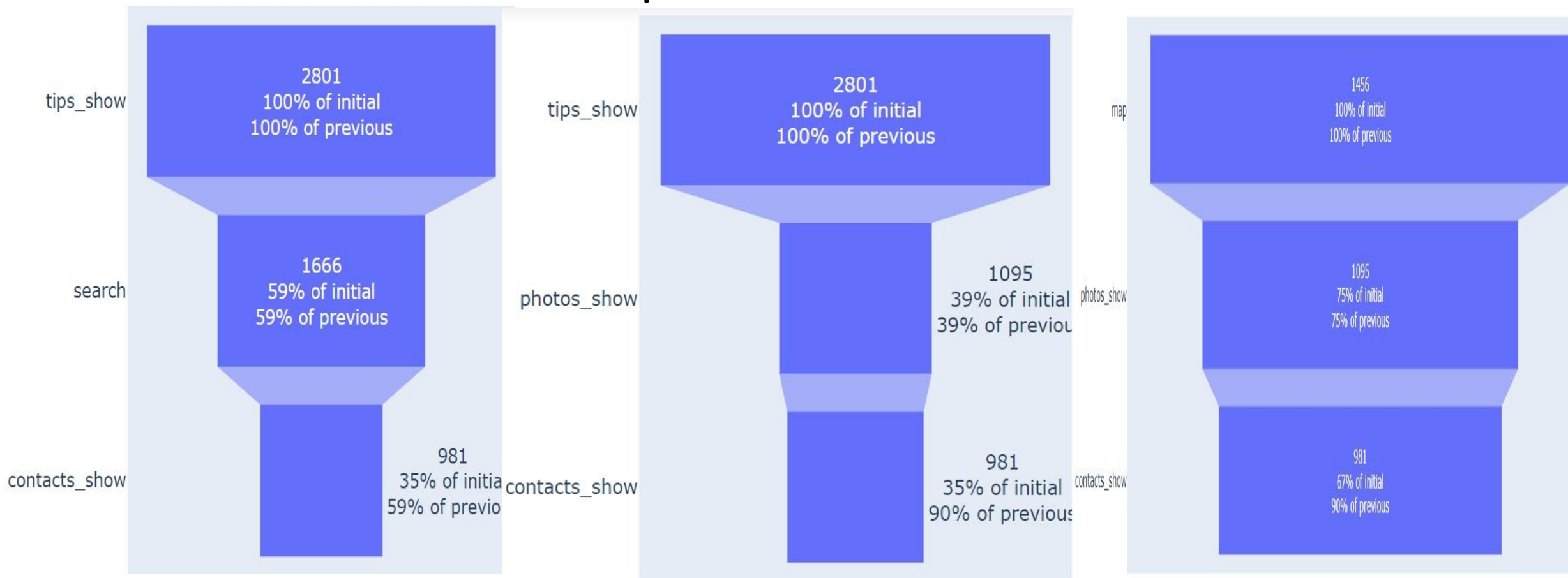


1. tips_show (2801 пользователь)

2. search(1666 пользователей) совершили наибольшие доли пользователей

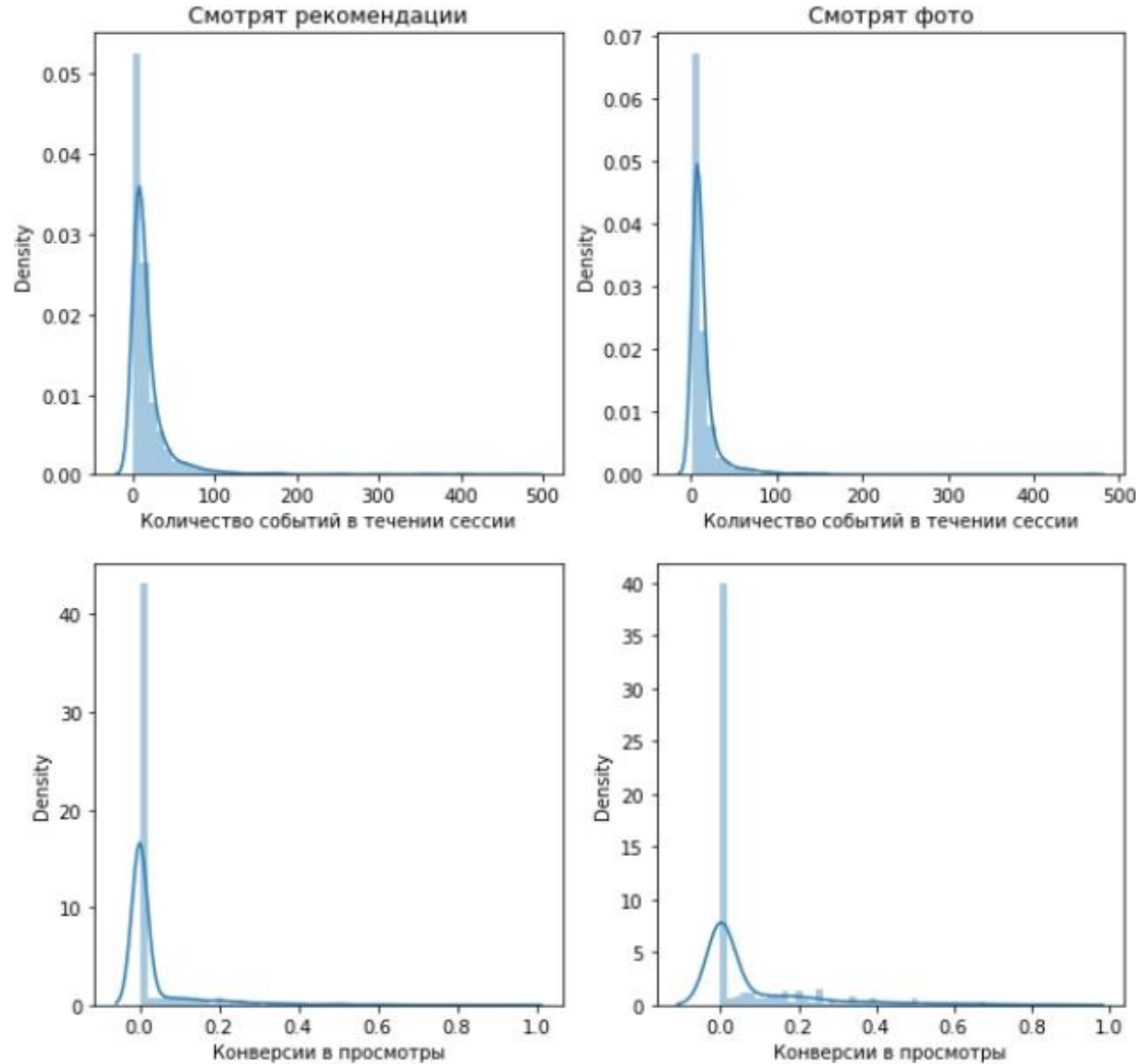
3е место попадает событие map, которое совершили почти 1,5 тыс. пользователей

Воронки событий



1. "классический" процент перехода к (ЦС) - 35% от начального, 59 % от предыдущего
 - 2-й вариант воронки "любителей фото" отличается высоким % перехода в ЦС после просмотра фото
 3. вариант воронки идеален и характеризует поведение целеустремленных пользователей
- Вариант 3 наиболее близок к реальности, но конверсия тоже не высока

Сегментация пользователей



НЕНУЖНЫЕ
ВЕЩИ

Группа 1

Часто пользуются и смотрят контакты.

- Активные пользователи, достигают цели

Работаем над другими группами, чтобы увеличить численность этой группы.

Группа 2

Активно пользуется приложением, но ни разу не смотрели контакты.

- Группа массовая по численности: активны, но не интересуются контактами.

Возможно, просто не могут найти то, что им нужно.

Нужно детальнее посмотреть, как они пользуются приложением: возможно, им недостаточно функционала приложения (нет возможности уточнения запроса, фильтров).

На данную группу следует фокусировать внимание.

Группа 3

Пользовались редко, но смотрели контакты.

- На эту группу также следует фокусировать внимание.

Они пользовались приложением лишь раз, но все равно достигли цели.

Можно попробовать "позвать" их обратно, используя уведомления и т.д.

Группа 4

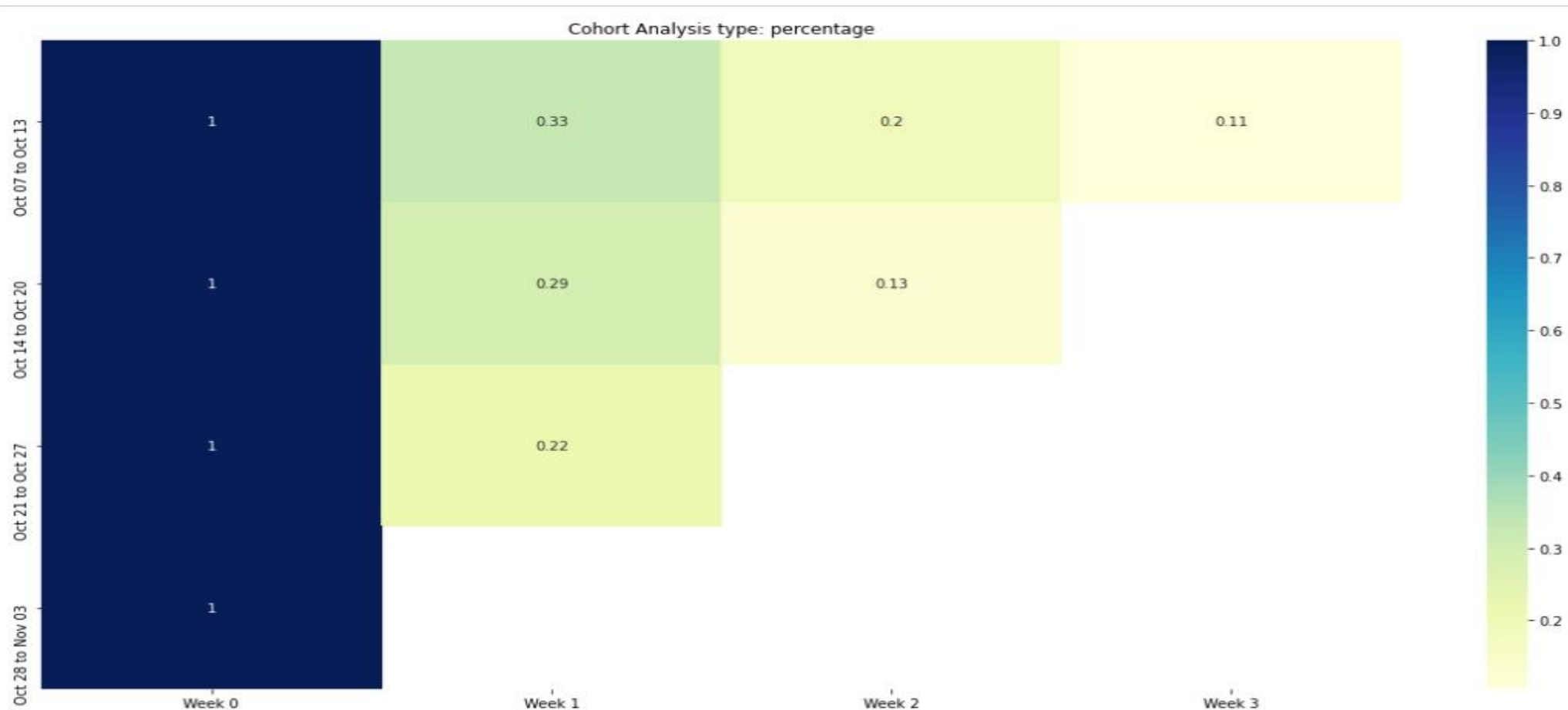
Почти не пользуются, контакты не смотрели ни разу.

- Самая значительная по численности группа, почти не пользуются и не достигают цели.

Необходимо искать способ "вернуть" их в приложение.

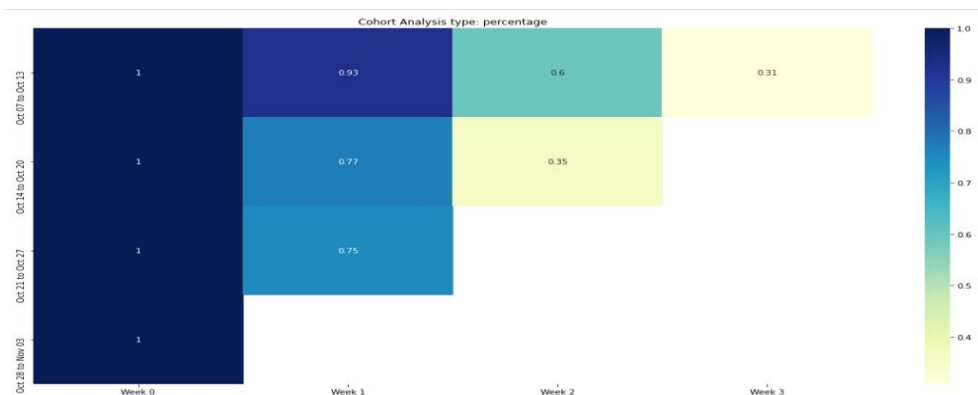
И. Кирилов

Ответы на вопросы заказчика

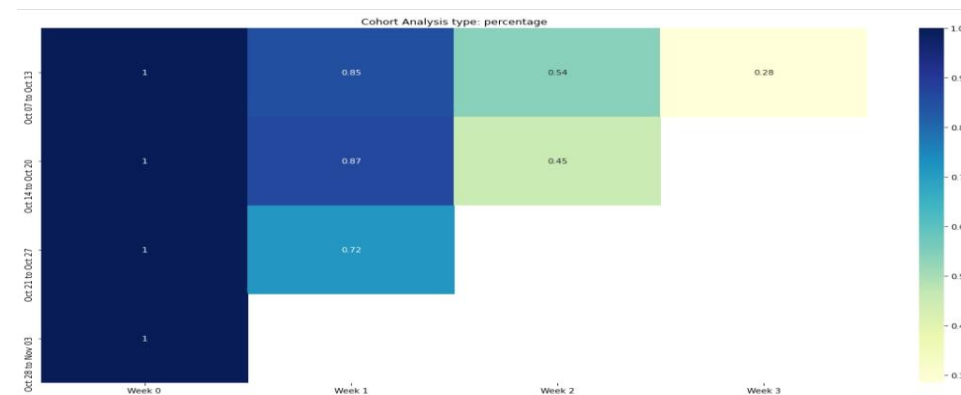


Retention rate снижается от недели к недели

Retention Rate



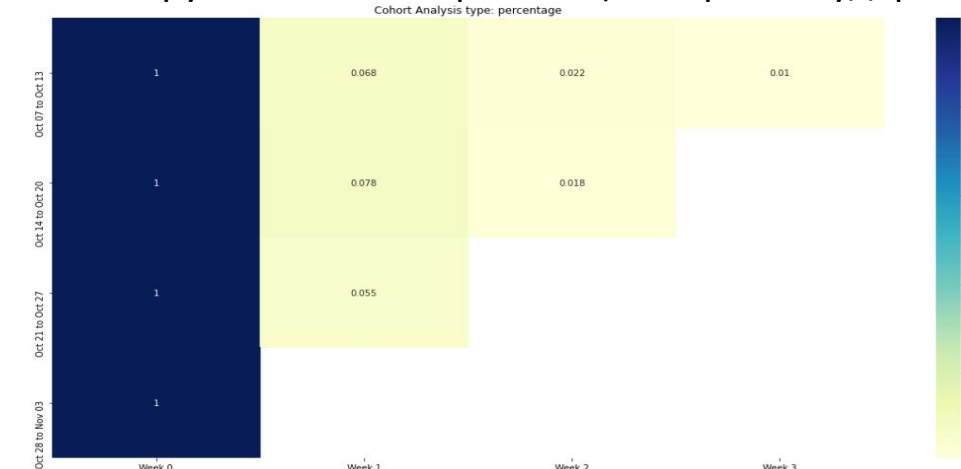
"лучшая" группа показывает хорошие результаты.



Активная группа без конверсии в ЦС - хорошее удержание



Не Активная группа с ЦС показала гораздо худшее Retention Rate по сравнению с верхними группами

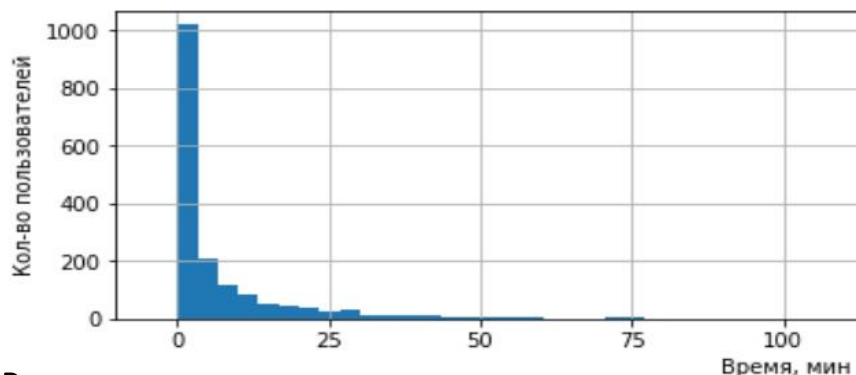


"худшая" группа - минимальная Retention Rate

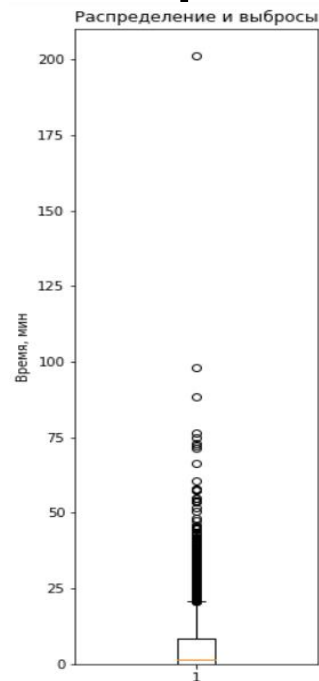
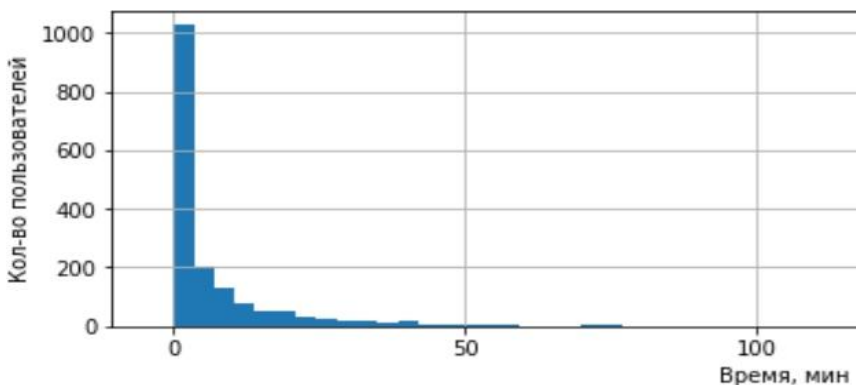
Группа пользователей (которые Пользовались редко, но смотрели контакты) демонстрирует максимальное количество ЦС (contacts_show).

Различие времени между распространенными событиями пользователей из разных групп

Распределение времени между распространенными событиями активных групп пользователей, совершивших ЦС



Распределение времени между распространенными событиями не активных групп пользователей, совершивших ЦС



5-й перцентиль	= 0.0 минут
10-й перцентиль	= 0.0 минут
15-й перцентиль	= 0.0 минут
20-й перцентиль	= 0.0 минут
25-й перцентиль	= 0.0 минут
30-й перцентиль	= 0.0 минут
35-й перцентиль	= 0.2 минут
40-й перцентиль	= 0.7 минут
45-й перцентиль	= 1.0 минут
50-й перцентиль	= 1.5 минут
55-й перцентиль	= 2.4 минут
60-й перцентиль	= 3.4 минут
65-й перцентиль	= 4.6 минут
70-й перцентиль	= 6.0 минут
75-й перцентиль	= 8.3 минут
80-й перцентиль	= 10.5 минут
85-й перцентиль	= 14.4 минут
90-й перцентиль	= 20.4 минут
95-й перцентиль	= 29.9 минут

Гистограммы, boxplot и перцентили показывают: конкретное значение усредненного показателя времени от первого события до целевого сложно определить, так как время довольно "плавно" возрастает от нескольких секунд до 29 минут.

Можно выделить:

85-й перцентиль = 14.4 минут - после него идет довольно резкий рост времени

90-й перцентиль = 20.4 минут

5 минут до 20 минут - рациональный диапазон времени

6 - 8 минут - можно принять как основной диапазон времени

Статистические гипотезы

Установка приложений из yandex и из google

Гипотеза 1:

Конверсии в просмотры контактов между теми пользователями, которые совершили установку приложения, прийдя с источника yandex и пользователя, совершившими установки, из источника google, различаются.

Формулировка нулевой гипотезы H_0 :

конверсии в просмотры контактов между теми пользователями, которые совершили установку приложения, прейдя с источника yandex и пользователя, совершившими установки, из источника google, равны.

Альтернативная гипотеза H_1 :

конверсии пользователей, совершивших установку из разных источников, различаются.

Доля пользователей группы a совершивших "contacts_show": 0.24357838795394154

Доля пользователей группы b совершивших "contacts_show": 0.24715615305067218

p-значение: 0.8244316027993777

Не получилось отвергнуть нулевую гипотезу, нет оснований считать доли разными

Вывод:

Статистической разницы конверсии в просмотры контактов между пользователями совершившими установку приложений с yandex и google - нет;

Источник установки никак не влияет на просмотры целевого события.

При увеличении действий разных групп пользователей конверсия в просмотры контактов различается

Гипотеза 2:

Конверсии в просмотры контактов между теми пользователями, которые совершили мало действий не различаются, а при увеличении действий различаются.

Формулировка нулевой гипотезы H_0 :

конверсии в просмотры контактов между теми пользователями, которые совершили мало действий равны.

Альтернативная гипотеза H_1 :

конверсии в просмотры контактов между теми пользователями, которые совершили много действий, различаются.

Гипотеза с порогом действий равным 2

p-значение: 0.03305937589743335

Не получилось отвергнуть нулевую гипотезу, нет оснований считать доли разными

Гипотеза с порогом действий равным 100

p-значение: 0.0

Отвергаем нулевую гипотезу: между долями есть значимая разница

Вывод:

Статистической разницы конверсии в просмотры контактов между пользователями, совершившими мало действий с порогом 2 - нет;

Статистическая разница конверсии в просмотры контактов между пользователями, совершившими много действий с порогом 100 - между долями есть значимая разница;

С увеличением действий значимость группы для исследований увеличивается.

РЕКОМЕНДАЦИИ:

- Уменьшить время поиска желаемого результата, быстрее предлагать конечный результат поиска, путем расширения фильтров при загрузке товара, с использованием качественного фото.
- Добавить возможность запроса по фото или картинке.
- Предложить пользователю вариант выбора поиска или товара (к примеру, на картинках, с вопросом: "Какая вещь больше нравится?", обратив действия в диалог)
- Обратить внимание на поведение самой многочисленной **Группы 4**, с целью развития приложения использовать визуальный характер информации.
Данная категория склонна к визуальному восприятию.

Спасибо за внимание!

[public.tableau](https://public.tableau.com)