

ПРОЕКТ

Мобильные приложения — Выделение групп пользователей на основе поведения

Выполнил: Габетдинов Фарид (farid.gabetdinov@yandex.ru)

Москва 2022

СОДЕРЖАНИЕ

1. ЦЕЛИ И ЗАДАЧИ ИССЛЕДОВАНИЯ
2. ОСНОВНОЙ ВЫВОД
3. ИСХОДНЫЕ ДАННЫЕ
4. ПОЛЬЗОВАТЕЛИ ПО ИСТОЧНИКАМ
5. ПОЛЬЗОВАТЕЛИ ПО ДНЯМ
6. СРЕДНИЕ ПОКАЗАТЕЛИ И ДИНАМИКА СОБЫТИЙ И СЕССИЙ
7. СОБЫТИЯ ПО ВИДАМ
8. ГРУППЫ ПОЛЬЗОВАТЕЛЕЙ. УДЕРЖАНИЕ. ЧАСТОТА ДЕЙСТВИЙ. ВРЕМЯ В ПРИЛОЖЕНИЕ И КОНВЕРСИЯ В ЦЕЛЕВОЕ
9. ПРОВЕРКА ГИПОТЕЗ
10. РЕКОМЕНДАЦИИ

Цели и задачи исследования

В рамках проекта необходимо было выделить группы пользователей, на основе событий выполняемых в мобильном приложении. Далее оценить их метрики:

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

Вопросы на которые необходимо ответить:

- Пользователи какой группы склонны часто возвращаться в мобильные приложения (RETENTION RATE)?
- Пользователи какой группы часто делают целевое событие (конверсия в целевое действие)?

В финале необходимо дать рекомендации по развитию приложения

Основной вывод

Пользователи у которых максимальная продолжительность сессии не превышает 4,5 минуты, проводят меньше всего времени в приложении, при этом имеют самые высокие показатели удержания и конверсии в целевое событие – просмотр контактов

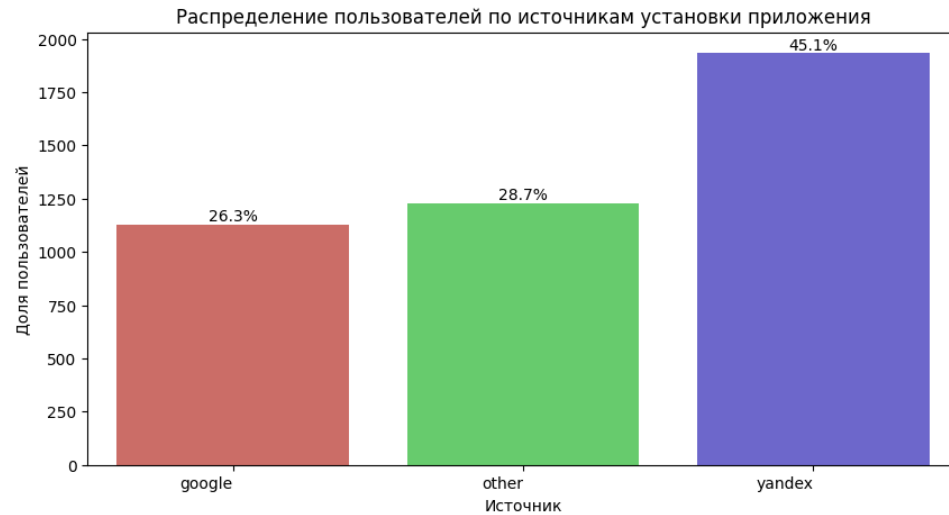
Исходные данные

На анализ были предоставлены два датасета:

- 1й - с данными об источниках, через которые пользователи устанавливали приложение,
- 2й - с данными об активностях этих же пользователей в приложении за период с 7 октября по 3 ноября 2019г.

В рамках предобработки оба датасета были объединены в один

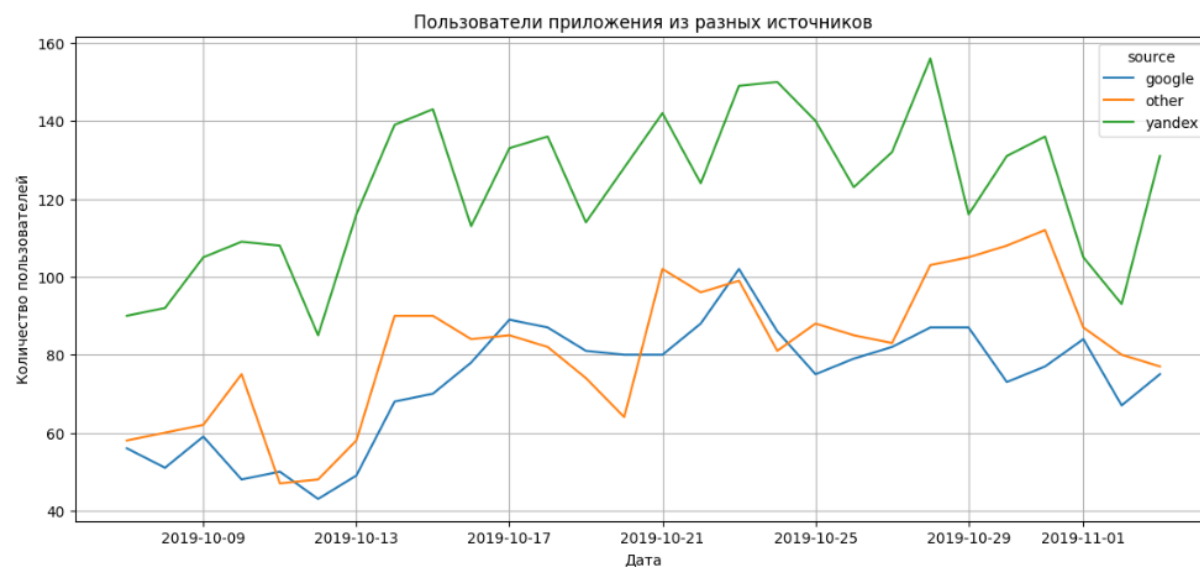
Пользователи. Источники



Источник	Кол-во пользователей	Доля
Yandex	1934	45%
Прочие	1230	29%
Google	1129	26%

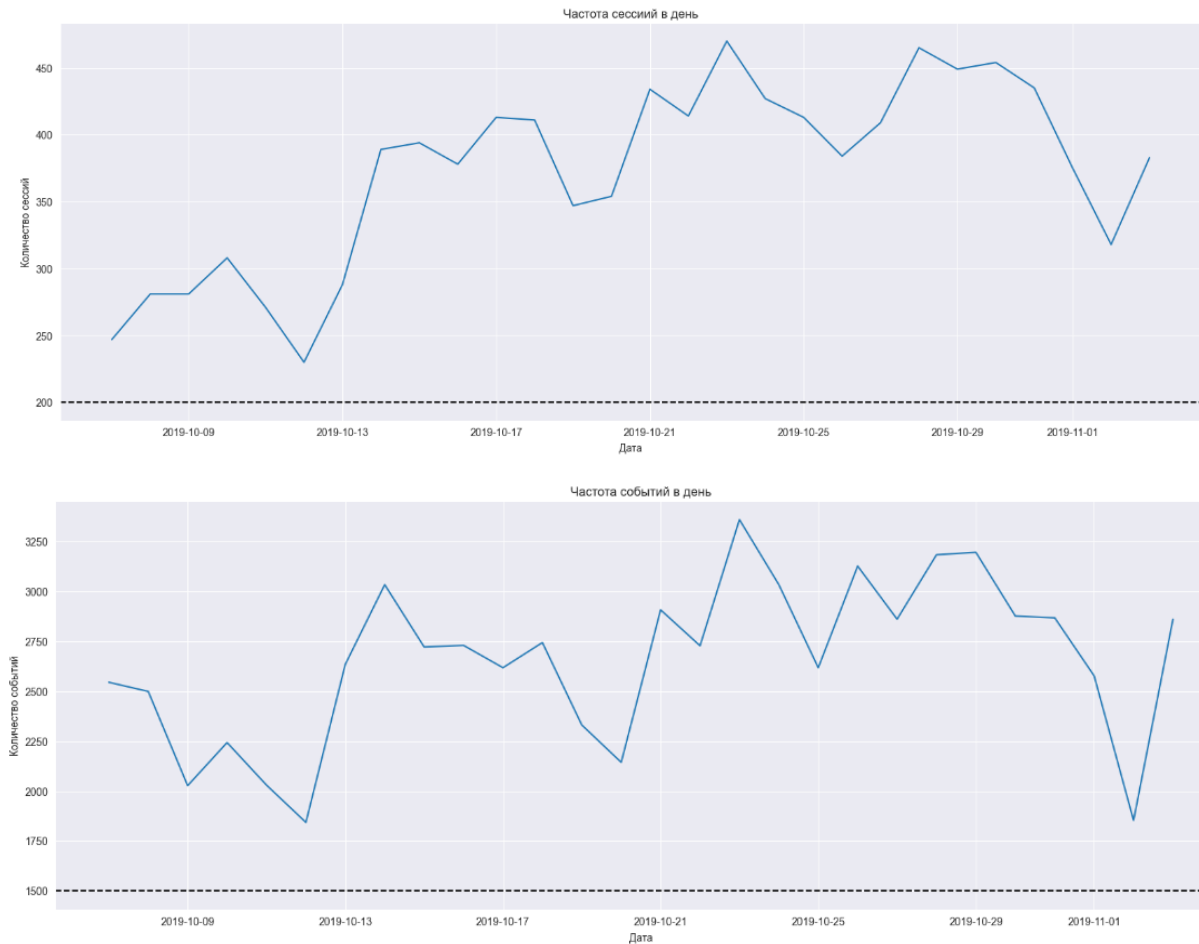
- За четыре недели (с 7 октября по 3 ноября 2019) приложением воспользовались 4293 новых пользователя
- Самое большое кол-во пользователей приложения появилось через канал Yandex - 1934 пользователя (45%). Это скорее всего объясняется тем, что приложение является локальным для пользователей России и стран СНГ.
- Далее идут источники «прочие» с 1230 пользователями (29%) и источник Google с 1129 пользователями (26%)
- В течении этого периода пользователи совершали от 1843 до 3360 действий ежедневно

Пользователи. Кол-во по дням и источникам



- кол-во новых Yandex пользователей всегда выше пользователей источников Other и Google. Источник Other, как правило, на втором месте (21 из 28 дней)
- В период с 7 по 28 октября наблюдается неравномерный рост числа новых пользователей у источника Yandex, с уровня 90 до 158. Далее идет спад до уровня 131
- У источника Other, наблюдается неравномерный рост числа новых пользователей с 59 до 111 с 7 по 31 октября, далее снижение до 79
- Источник Google растет по новым пользователям с 7 по 23 октября, с 58 до 101 нового пользователя. Далее снижение до уровня 75 3 ноября
- Чаще всего пользователи приходили из Yandex, примерно одинаково из Google и других источников.

Средние показатели событий и сессий. Динамика сессий и событий



- В целях построения пользовательских профилей, их сегментации и сравнительного анализа метрик, было выделено 10368 уникальных сессий совершенных уникальными пользователями.
- 21% сессий получен с «нулевой» продолжительностью – это пользователи, которые просто убивали время за просмотром объявлений и фотографий размещенных в приложении не особо углубляясь в их содержание
- Максимальное количество сессий на одного пользователя не превышает 100 сессий
- Средняя продолжительность сессии составляет не более 6 мин
- Было посчитано, что в среднем (по медиане) на одного уникального пользователя приходится 9 событий при одном среднем количестве сессий, среднее кол-во событий в сессии равно 4. При этом, эти показатели не зависят от источника установки приложения.
- Активное взаимодействие пользователей с приложением происходит в течении рабочей недели, а в выходные дни активность падает. Также видно, что, как частота сессий, так и частота событий растут в течении первых трех недель рассматриваемого периода - возможно это результат действия промо акций на начало периода, о которых нам не известно. Далее, в течении недели наблюдается так называемое плато (по крайней мере по сессиям), которое, на 29 октября заканчивается снижением до 2 ноября

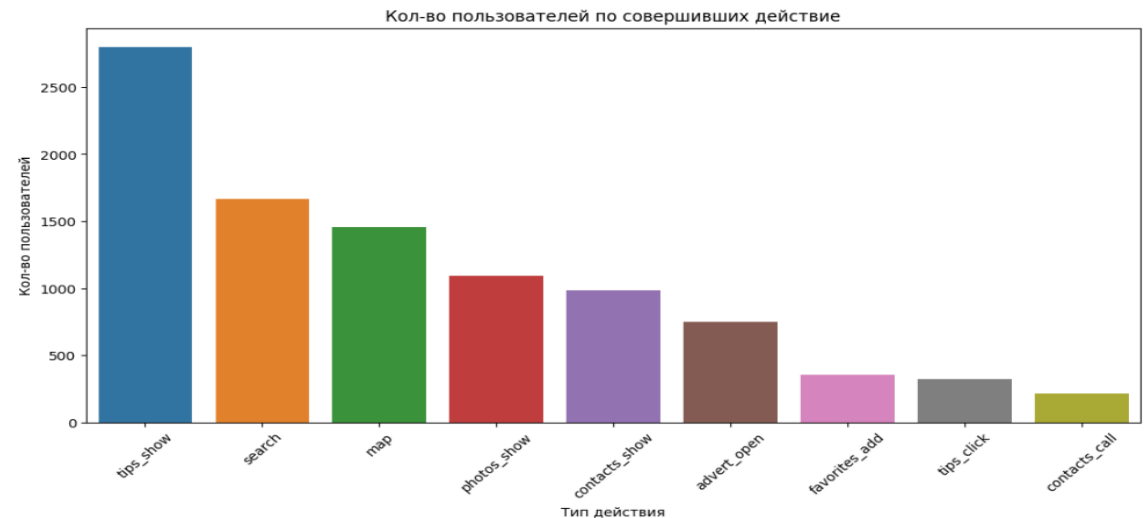
События по видам

Событие	Кол-во пользователей совершивших событие	Доля пользователей совершивших событие, %
tips_show	2801	65,25
search	1666	38,81
map	1456	33,92
photos_show	1095	25,51
contacts_show	981	22,85
advert_open	751	17,49
favorites_add	351	8,18
tips_click	322	7,5
contacts_call	213	4,96

Для того чтобы построить продуктовую воронку нужно определить в какой последовательности происходят события. Из условий и документации нельзя точно определить последовательность событий, этот момент нужно уточнять у разработчиков и только тогда можно будет построить правильную воронку.

По количеству уникальных пользователей совершивших события можно сказать следующее:

- самое большое кол-во пользователей совершили действие tips_show, т.е. пользователь увидел рекомендованные объявления
- поиск совершает 38.81%
- открытие карты совершает 33.92%
- целевое событие contacts_show совершает 22.85% пользователей от общего количества.

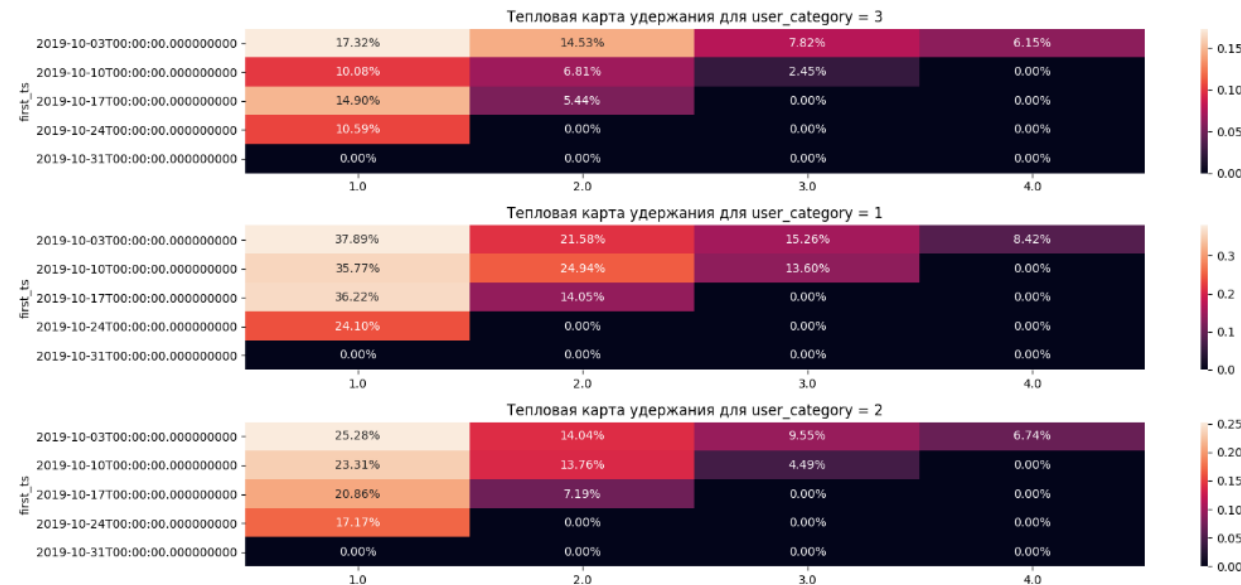


Группы пользователей. Удержание

Группа	Кол-во пользователей	Длительность сессии
1	1416	от 0 до 4,5 мин
2	1477	от 4,5 до 14 мин
3	1400	от 14 мин и выше

- Определим группы следующим образом:
- группа 1 - длительность сессии от 0 до 4,5 минут
- группа 2 - длительность сессии от 4,5 до 14 минут
- группа 3 - длительность сессии больше 14 минут

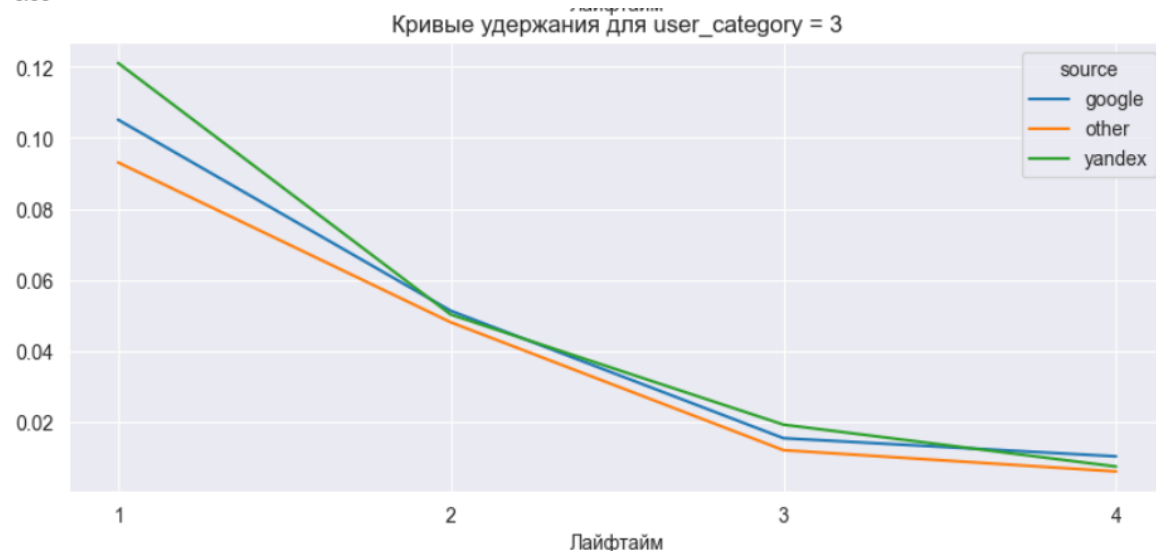
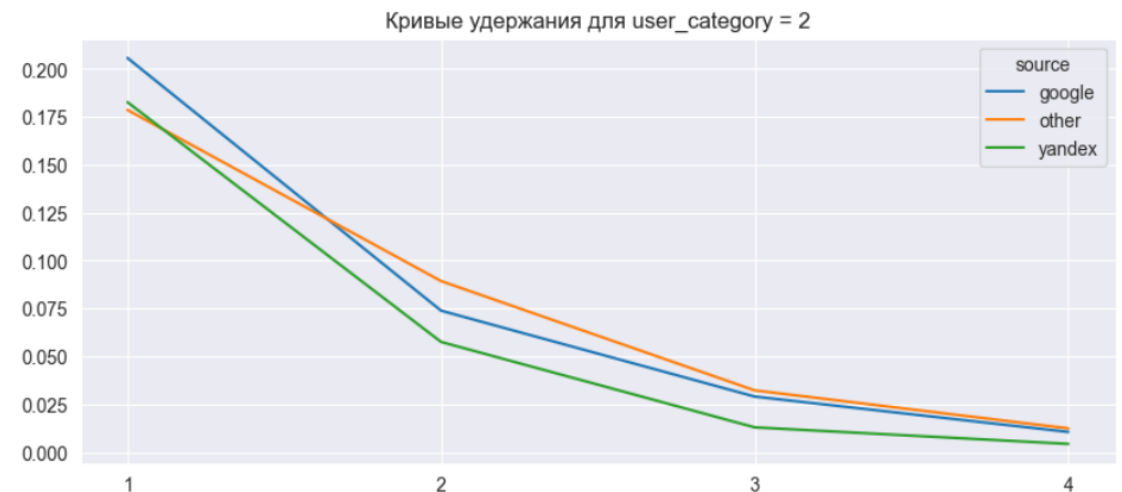
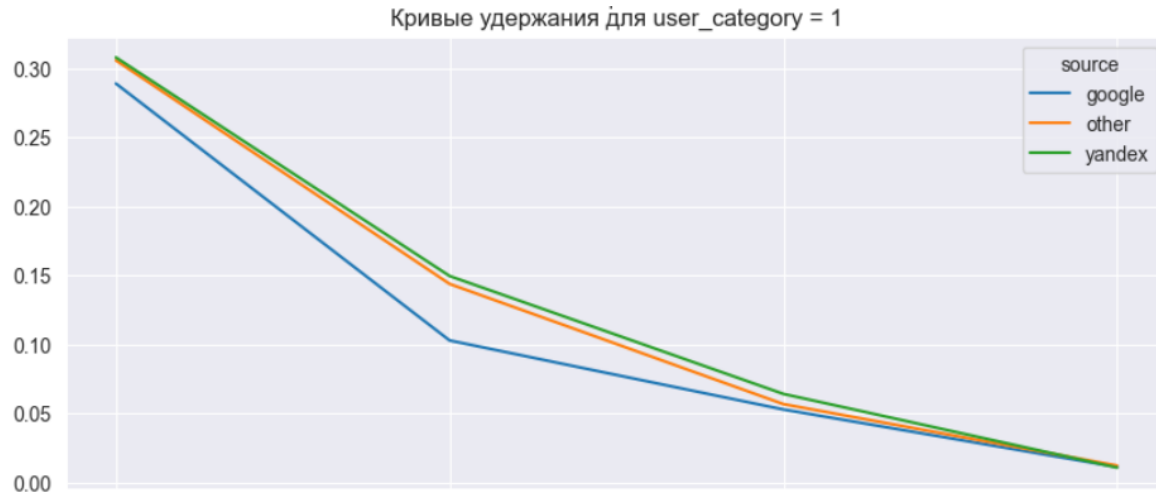
Данная группировка по длительности сессии обеспечивает более равномерное распределение пользователей между группами



Удержание

- для всех категорий в первую неделю удержание хуже у привлеченных с 24 по 31 число
- хуже всего удержание у 3й категории пользователей
- лучше всего удержание у 1й категории

Группы пользователей. Удержание (продолжение)



- по все группам, и это очевидно, удержание к концу 4й недели падает
- по 1-ой группе, самое лучшее удержание у пользователей Yandex, далее идут прочие источники и хуже всего у пользователей с Google
- по 2-ой группе, лучше удержание у "прочих" пользователей, на втором месте удержание пользователей с Google, и на третьем - Yandex пользователей
- по 3-ей группе, удержание Yandex пользователей на первом месте, на втором - удержание пользователей Google, с других ресурсов удержание заметно хуже

Группы пользователей. Частота действий

Группа	Кол-во событий
1	16702
2	21951
3	35544



Tips_show (пользователь увидел рекомендованные объявления). Чаще всего это событие выполняется пользователями 3 гр., далее 2я и первая

- Photos_show (просмотр фотографий): чаще всех смотрят пользователи 1й и 2й гр.
- Search (все виды поиска): 1й и 2й группы на первом и втором местах
- Advert_open (открытие карточки объявления): 2 и 3я гр чаще выполняют это событие чем 1я гр
- Contacts_show (просмотр контактов): целевое событие чаще всех выполняют пользователи 1й группы, затем 3й и 2й
- Map (пользователь открыл карту): открывают карту объявлений чаще пользователи 2й группы
- Favorites_add (добавление объявления в избранное): пользователи 1й группы выполняют чаще всех

Группы пользователей. Время в приложении и конверсия в целевое действие

Группа	Время в приложении	Конверсия в целевое событие, %
1	15 дн. 02:34:07	25,00
2	24 дн. 05:03:11	21,39
3	53 дн. 09:39:54	22,21

Меньше всего времени в приложении проводили пользователи первой группы. Учитывая, что самая высокая конверсия также принадлежит первой группе можно сказать, что у них самый позитивный пользовательский опыт. Т.е. по совокупности факторов: у них ясное представление о том что они ищут, они это как правило быстро находят и, имея положительный опыт (retention rate) у них есть желание вернуться в приложение.

Проверка гипотез.

Гипотеза №1

Основная гипотеза: конверсия в просмотры контактов пользователей пришедших через Yandex не отличается от конверсии пользователей пришедших через

В результате теста сделан вывод:

В результате проверки статистической гипотезы о равенстве долей z-критерием оказалось, что разница в конверсии между пользователями, пришедшими из поисковика Yandex, и теми, кто пришёл из Google, не является статистически значимой. Гипотезу H_0 отвергнуть не получилось

Гипотеза №2

Основная гипотеза: конверсия в просмотры контактов у пользователей, которые совершают действия tips_show и tips_click, не отличается от конверсии у пользователей, которые совершают только действие tips_show.

В результате теста сделан вывод:

В результате проверки статистической гипотезы о равенстве долей z-критерием оказалось, что разница в конверсии между пользователями совершающими действия tips_show+tips_click, и теми, совершает только tips_show, является статистически значимой и отвергаем гипотезу H_0

Рекомендации:

По результатам проведенного исследования можно предложить следующие рекомендации по развитию приложения:

- приложить усилия по сокращению времени пользовательских сессий и времени проведенного в приложении, это даст рост RR и CR. Это все, что связано пользовательским опытом: грамотная структура приложения, интуитивно удобная навигация. Например, просмотр фотографий — самое частое пользовательское событие, после того как пользователь увидел рекламное объявление. В этой связи следует мотивировать пользователей размещать больше фотографий вещей, это способствует продажам и соответственно росту вовлеченности.
- самое большое кол-во привлеченных пользователей было получено через источник Yandex. В этой связи следует размещать таргетированную рекламу пользователям экосистемы Yandex. Не стоит оставлять без внимания и другие источники: категория "прочие" и Google. Попробовать найти точки роста по привлечению новых пользователей из этих источников.
- в виду отсутствия понимания алгоритма поведения пользователей, нет возможности построить воронку событий и с этой позиции проанализировать динамику RR и CR. В этой связи необходима работа по определению и анализу алгоритма взаимодействия пользователя с приложением
- если есть такая возможность, то дополнить данные информацией, допустим регион из которого было взаимодействие с приложением, с какого девайса: Android, iPhone, PC, Mac, "логинились".

СПАСИБО!