

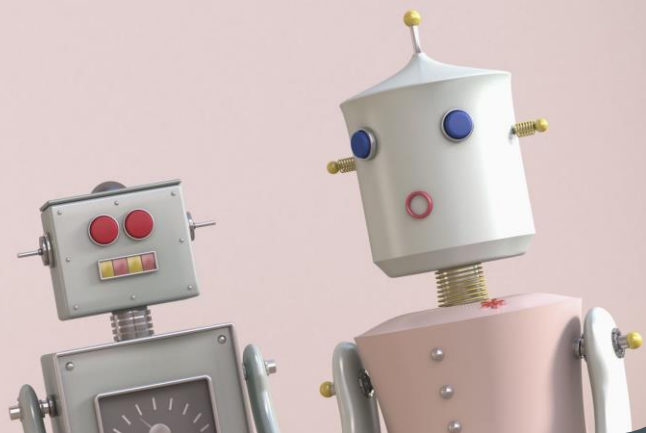


ФИНАЛЬНЫЙ ПРОЕКТ



АНАЛИЗ ДАННЫХ





ОПИСАНИЕ ПРОЕКТА

- Мы решили подготовить анализ поведения пользователей мобильной игры «Космические братья»;
- Цели:
 - Анализ влияние источника перехода в игру на поведение пользователя;
 - Проверка статистических гипотез:
 - Время завершения уровня различается в зависимости способа прохождения;
 - Количество дней до завершения первого уровня различается в зависимости от источника трафика



ДАННЫЕ (1/2)

	event_datetime	event	building_type	user_id	project_type
135635	2020-06-05 00:08:06	building	research_center	f21d179f-1c4b-437e-b9c6-ab1976907195	NaN
135636	2020-06-05 02:25:12	finished_stage_1	NaN	515c1952-99aa-4bca-a7ea-d0449eb5385a	NaN
135637	2020-06-05 08:57:52	building	research_center	ed3e7d02-8a96-4be7-9998-e9813ff9c316	NaN
135638	2020-06-05 12:12:27	finished_stage_1	NaN	32572adb-900f-4b5d-a453-1eb1e6d88d8b	NaN
135639	2020-06-05 12:32:49	finished_stage_1	NaN	f21d179f-1c4b-437e-b9c6-ab1976907195	NaN

В датасете содержатся данные первых пользователей приложения — когорты пользователей, которые начали пользоваться приложением в период с 4 по 10 мая включительно.

Датасет game_actions.csv:

- `event_datetime` — время события;
- - `event` — одно из трёх событий:
 1. `building` — объект построен,
 2. `finished_stage_1` — первый уровень завершён,
 3. `project` — проект завершён;
- - `building_type` — один из трёх типов здания:
 1. `assembly_shop` — сборочный цех,
 2. `spaceport` — космопорт,
 3. `research_center` — исследовательский центр;
- - `user_id` — идентификатор пользователя;
- - `project_type` — тип реализованного проекта;



ДАННЫЕ (2/2)

	source	day	cost
0	facebook_ads	2020-05-03	935.9
1	facebook_ads	2020-05-04	548.4
2	facebook_ads	2020-05-05	260.2
3	facebook_ads	2020-05-06	178.0
4	facebook_ads	2020-05-07	111.8

	user_id	source
0	0001f83c-c6ac-4621-b7f0-8a28b283ac30	facebook_ads
1	00151b4f-ba38-44a8-a650-d7cf130a0105	yandex_direct
2	001aaea6-3d14-43f1-8ca8-7f48820f17aa	youtube_channel_reklama
3	001d39dc-366c-4021-9604-6a3b9ff01e25	instagram_new_adverts
4	002f508f-67b6-479f-814b-b05f00d4e995	facebook_ads

Датасет ad_cost.csv:

- - `day` - день, в который был совершен клик по объявлению
- - `source` - источник трафика
- - `cost` - стоимость кликов

Датасет user_source.csv содержит колонки:

- - `user_id` - идентификатор пользователя
- - `source` - источников, с которого пришёл пользователь, установивший приложение



ОСНОВНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ

- Маркетинговая кампания проводилась с 03.05.2020 по 09.05.2020.

- Ее период выпадает на майские праздники, когда люди обычно на отдыхе и могут меньше играть в игры



Источники

- С Яндекс пришла наибольшая доля игроков. На втором месте Instagram, далее идут Facebook и Youtube.
- ЦА кампании русскоговорящая из-за большого трафика с Яндекса.
- Минимальные САС у Youtube — 0,4, что довольно странно, а самый дорогой у Facebook, что ожидаемо, если наша ЦА молодые люди.
- Необходимо скорректировать маркетинговую стратегию продвижения, уменьшив вливания в Facebook и перекинуть их на Яндекс.
- Гипотеза, что источник трафика не зависит от способа завершения первого уровня подтверждена.



Кол-во пользователей

- Среднее количество уникальных пользователей за исследуемый период:
 - в день: 2 884
 - в неделю: 6 120
- После 12 дня в когортах активны не более 20% первоначального количества игроков.
- Около половины игроков завершили уровень 1:
- Пользователи, которые не завершили первый уровень, ушли из игры до того как сделают 7 построек.
- Нужно сделать deep dive поведения пользователей на этом этапе и возможно ввести мотивирующие действия в игре для предотвращения ухода.



ОСНОВНЫЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ



Время в игре

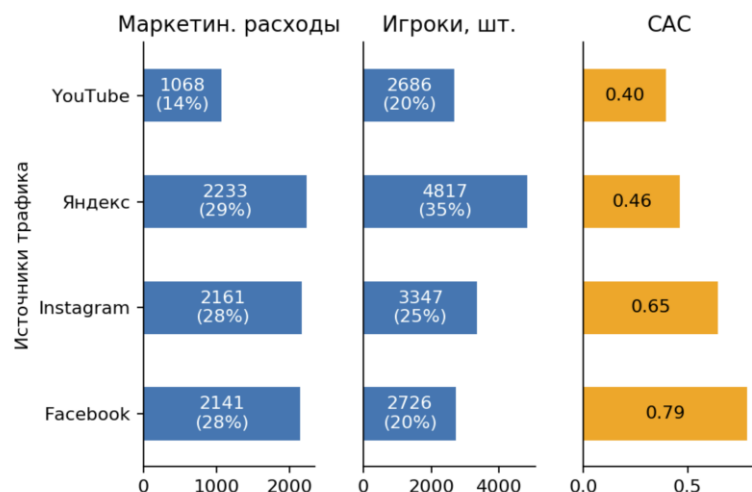
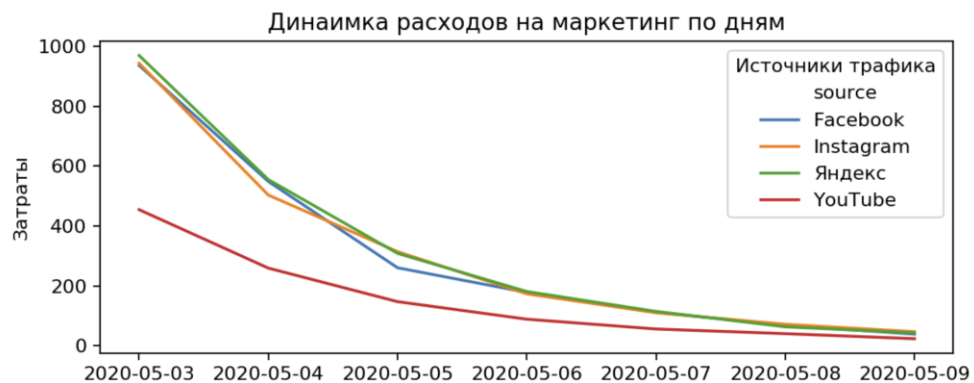
- Среднее время до завершения уровня 1:
 - Победа над врагом, дни: 12.
 - Реализация проекта, дни: 12.
- Гипотеза о равенстве времени завершения уровня 1 в зависимости от способа прохождения не подтверждена. Отвергаем нулевую гипотезу. Время завершения уровня зависит от способа прохождения.
- Стандартное отклонение времени на завершение уровня 1 по условию победы больше, чем через постройку проекта. Вероятно, слабым игрокам попадаются слишком сильные противники и наоборот.
- Рекомендовал был провести анализ балансировки и подбора противников по уровню прокаченности



Постройки в игре

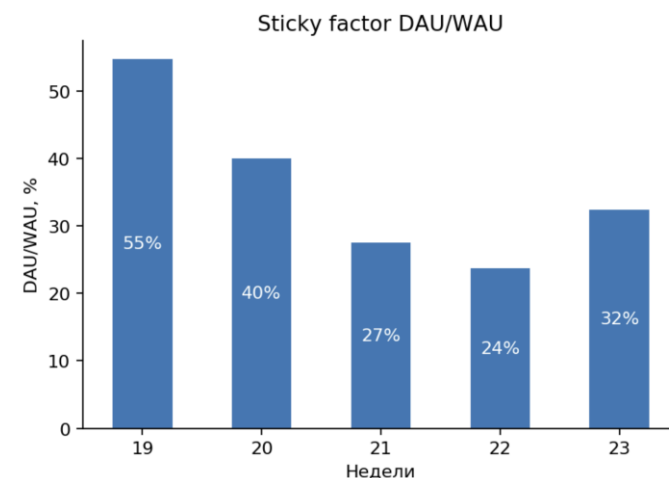
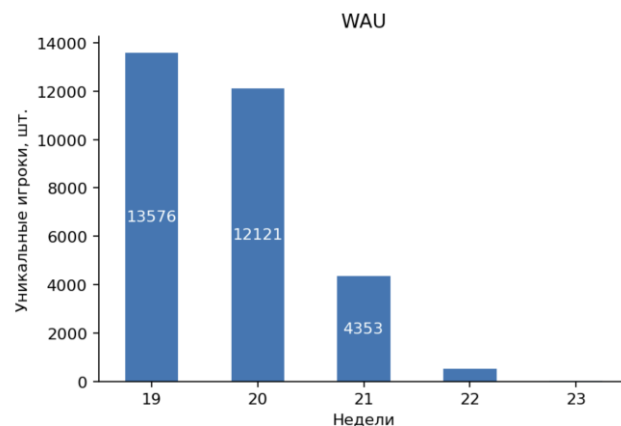
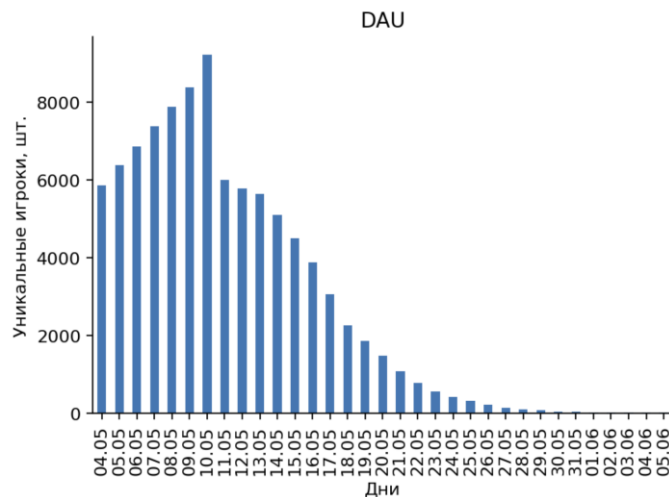
- Среднее количество построек до завершения первого уровня:
 - Для выполнивших первую победу над врагом: 10 шт.
 - Для тех, чтобы Реализация проекта: 13 шт.
- Вероятно, что баланс построек на завершение уровня 1 не выдержан. Игроки смогут обойти систему прокачки и покупки зданий через победу над врагом.
- Необходимо сбалансировать игровую систему до введения монетизации.

МАРКЕТИНГОВЫЕ ПОКАЗАТЕЛИ



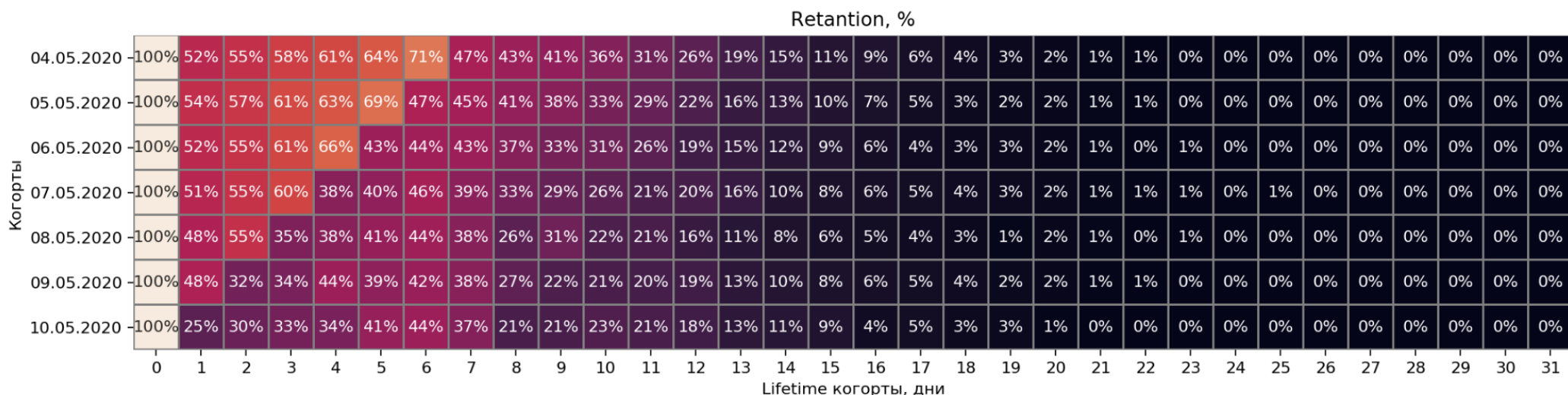
- Бюджет распределен пропорционально за исключением YouTube, на который потрачено меньше всего. Видимо нашей целевой аудитории там нет.
- С Яндекса пришла больше всего игроков, наверное целевая аудитория скорее всего русскоязычная. После идет Instagram и на третьем месте Facebook и Youtube с примерно одинаковым количеством игроков
- Наименьший показатель CAC по Youtube, что очень странно, обычно это самый дорогой канал в игровой индустрии с маленькой конверсией и как следствие высоким CAC
- Так как ЦА у нас русскоязычная, то ожидаемо, что Facebook будет самым дорогим каналом. Плюс скорее всего по возрасту наша аудитория может быть ближе к ВК, чем FB. Это может быть причина высокого CAC по данному источнику.

2 АНАЛИЗ КОЛИЧЕСТВА ПОЛЬЗОВАТЕЛЕЙ (MAU, DAU, КОГОРНЫЙ АНАЛИЗ И LIFETIME) (1/2)



- Все пользователи, которые установили игру, совершили хотя бы одно действие. Скорее всего у нас выгрузка только по пользователям, которые совершили хоть одно действие, а не по всем пользователям, которые скачали игры, т.к. слишком классные показатели.
- Первую неделю рост пользовательской базы обеспечивает маркетинговая кампания. Пиковое значение 10 тыс. пользователей в день приходится на 10 день - это последний день кампании.
- Пик снижения пользователей происходит на третью неделю
- Также себя ведет и динамика Sticky factor DAU/WAU. Небольшое увеличение в конце связано с колебаниями небольших абсолютных значений.

2 АНАЛИЗ КОЛИЧЕСТВА ПОЛЬЗОВАТЕЛЕЙ (MAU, DAU, КОГОРНЫЙ АНАЛИЗ И LIFETIME) (2/2)

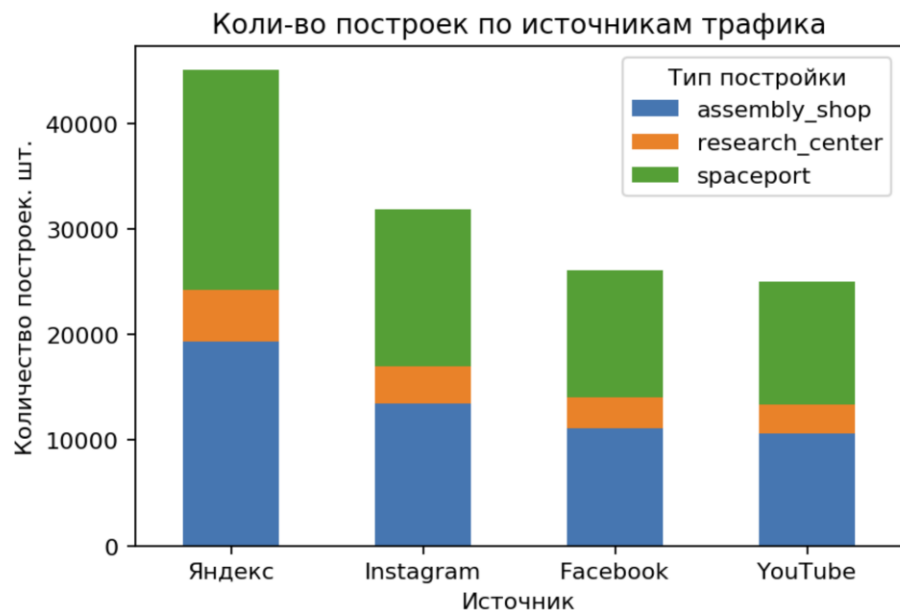


- В первые 6 дней жизни уровень удержания растет: Вероятно, что отчасти это связано, с проведением маркетинговой кампании. Посмотрим это ниже
- Через 12 дней уровень удержания снижается до менее 20%

3

АНАЛИЗ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ (ПО КЛЮЧЕВЫМ ДЕЙТСВИЯМ) В ЗАВИСИМОСТИ ОТ ИСТОЧНИКА ТРАФИКА

КОЛ-ВО ПОСТРОЕК ПО ИСТОЧНИКАМ ТРАФИКА

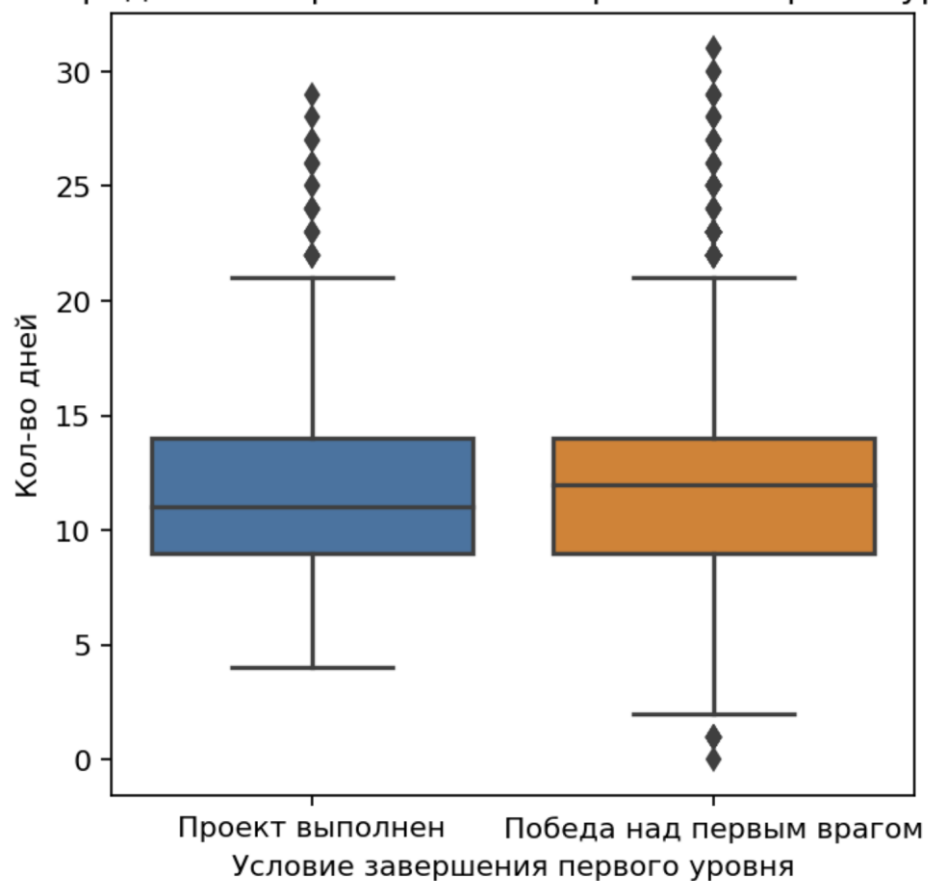


- Исходя из графика можно сделать предположение, что пользователи, которые не завершили первый уровень, ушли из игры после 6 постройки.
- Нужно сделать deep dive поведения пользователей на этом этапе и возможно ввести мотивирующие действия в игре для предотвращения ухода

4

АНАЛИЗ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ (ПО КЛЮЧЕВЫМ ДЕЙТСВИЯМ) - ВРЕМЯ НА ЗАВЕРШЕНИЕ УРОВНЯ 1

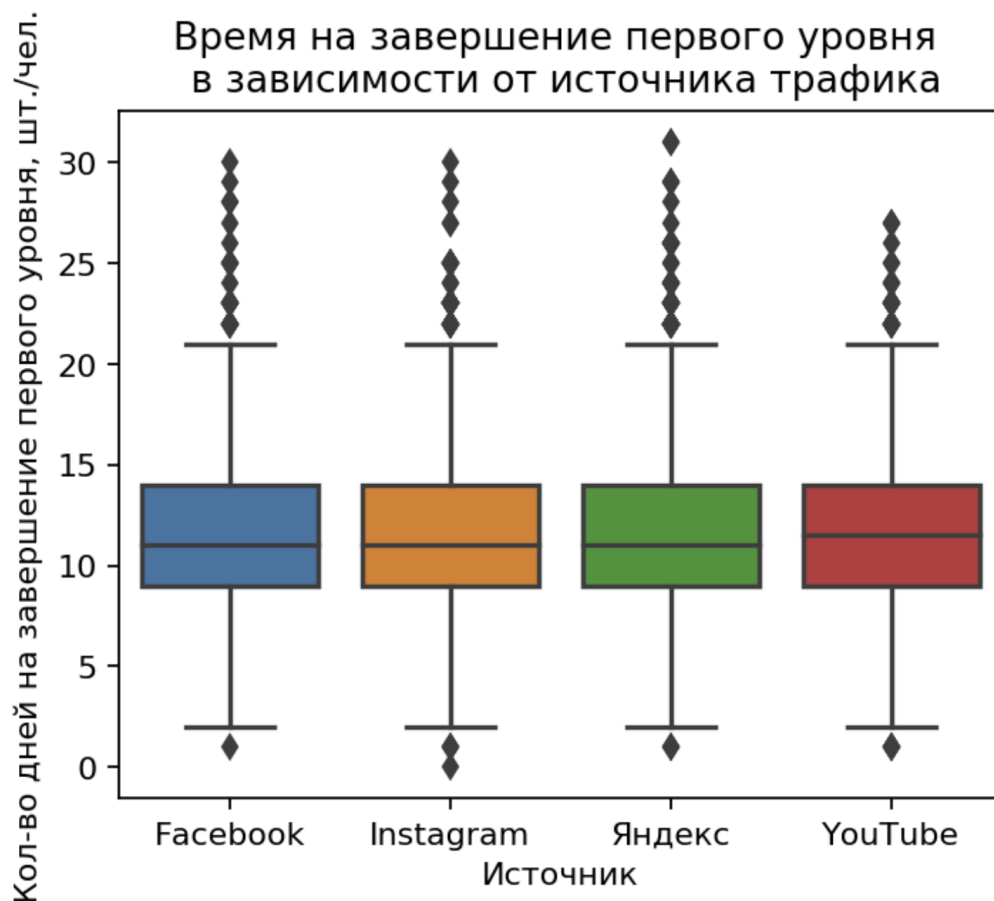
Распределение времени на завершение первого уровня



- В среднем, пользователи завершают первый уровень за 11,8 дней независимо от условия завершения.
- Медианное значение по кол-ву дней для завершения первого уровня за счет проекта составляет 11 дней, а за счет первой победы 12 дней.
- Стандартное отклонение времени на завершение уровня 1 по условию победы больше, чем через постройку проекта. Возможно, что нет баланса между слабыми и сильными противниками.

5

АНАЛИЗ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ (ПО КЛЮЧЕВЫМ ДЕЙТСВИЯМ) - ВРЕМЯ НА ЗАВЕРШЕНИЕ ПЕРВОГО УРОВНЯ В ЗАВИСИМОСТИ ОТ ИСТОЧНИКА ТРАФИКА



- Данный график очень наглядный и говорит сам за себя буквально. Нет каких-либо видимых отличий в значении медианы
- $1,5 \cdot Q$ для всех визуально одинаковый
- Отличие в данных относится к выбросам, что является нормальным

6 ПРОВЕРКА ГИПОТЕЗЫ: ВРЕМЯ ЗАВЕРШЕНИЯ УРОВНЯ РАЗЛИЧАЕТСЯ В ЗАВИСИМОСТИ СПОСОБА ПРОХОЖДЕНИЯ

- Различие времени завершения уровня в зависимости способа прохождения через реализацию проекта или через победу над первым игроком должно быть проверено через гипотезу о равенстве средних двух генеральных совокупностей:
- **H₀**: время завершения уровня не зависит от способа прохождения (среднее время через реализацию проекта равно среднему времени через победу над первым игроком).
- **H₁**: время завершения уровня зависит от способа прохождения.
- Среднее время завершения первого уровня:
 - - Выполнение проекта, дни: 11.8.
 - - Первая победа над врагом, дни: 11.8.
- р-значение: 0.846
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от способа прохождения.

7

ПРОВЕРКА ГИПОТЕЗЫ: СОБСТВЕННАЯ ГИПОТЕЗА (КОЛ-ВО ДНЕЙ ДО ЗАВЕРШЕНИЯ ПЕРВОГО УРОВНЯ РАЗЛИЧАЕТСЯ В ЗАВИСИМОСТИ ОТ ИСТОЧНИКА ТРАФИКА) (1/2)

- Кол-во дней до завершения первого уровня в зависимости от источника трафика должно быть проверено через гипотезу о равенстве средних двух генеральных совокупностей:
- **H₀**: количество дней до завершения первого уровня не зависит от источника трафика (среднее кол-во дней по источнику_1 равно среднему кол-во по источнику_2).
- **H₁**: количество дней до завершения первого уровня зависит от источника трафика.

- Проверка равенства ср. времени прохождения уровня 1 пользователям и, пришедших из Instagram и Facebook
- Среднее время завершения первого уровня:
 - - Пользователей из источника Instagram, дни: 11.9.
 - - Пользователей из источника Facebook, дни: 11.8.
- р-значение: 0.679
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.
- =====
- Проверка равенства ср. времени прохождения уровня 1 пользователям и, пришедших из Instagram и Яндекс
- Среднее время завершения первого уровня:
 - - Пользователей из источника Instagram, дни: 11.9.
 - - Пользователей из источника Яндекс, дни: 11.7.
- р-значение: 0.271
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.

7

ПРОВЕРКА ГИПОТЕЗЫ: СОБСТВЕННАЯ ГИПОТЕЗА (КОЛ-ВО ДНЕЙ ДО ЗАВЕРШЕНИЯ ПЕРВОГО УРОВНЯ РАЗЛИЧАЕТСЯ В ЗАВИСИМОСТИ ОТ ИСТОЧНИКА ТРАФИКА) (2/2)

- **Проверка равенства ср. времени прохождения уровня 1 пользователя ми, пришедших из Instagram и YouTube**

- Среднее время завершения первого уровня:
 - - Пользователей из источника Instagram, дни: 11.9.
 - - Пользователей из источника YouTube, дни: 11.9.
- р-значение: 0.738
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.

=====

- **Проверка равенства ср. времени прохождения уровня 1 пользователя ми, пришедших из Facebook и Яндекс**

- Среднее время завершения первого уровня:
 - - Пользователей из источника Facebook, дни: 11.8.
 - - Пользователей из источника Яндекс, дни: 11.7.
- р-значение: 0.560
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.

- **Проверка равенства ср. времени прохождения уровня 1 пользователями, пришедших из Facebook и YouTube**

- Среднее время завершения первого уровня:
 - - Пользователей из источника Facebook, дни: 11.8.
 - - Пользователей из источника YouTube, дни: 11.9.
- р-значение: 0.483
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.

=====

- **Проверка равенства ср. времени прохождения уровня 1 пользователями, пришедших из Яндекс и YouTube**

- Среднее время завершения первого уровня:
 - - Пользователей из источника Яндекс, дни: 11.7.
 - - Пользователей из источника YouTube, дни: 11.9.
- р-значение: 0.172
- Не получилось отвергнуть нулевую гипотезу. Время завершения первого уровня не зависит от источника трафика.