

Тестовое задание в [REDACTED]
на позицию аналитика данных
Бересневой-Батт Валерии

Отчет по анализу результатов A/B-теста мобильной игры [REDACTED]

Задача: проанализировать, как на метриках проекта сказалось уменьшение скорости восстановления жизни с 30 до 3 минут.

Описание данных: смоделированные данные пользователей за первых 7 дней жизни в игре.

План проекта:

1. загрузить данные, сделать их предобработку и декомпозицию;
2. проверить правильность распределения пользователей по группам;
3. проанализировать прохождение пользователями уровней (старты, финиши и их виды) в контрольной и тестовой группах;
4. проанализировать получение и траты золота в обеих группах;
5. посмотреть на изменение метрик Retention rate, Churn rate, Conversion rate и LTV в группах;
6. проверить гипотезы;
7. сформулировать выводы об успешности проведения теста.

Примечание:

- *В данных отсутствуют даты прихода в приложение и совершения действий, поэтому когорты будут определены только по принадлежности к контрольной или тестовой группе.*
- *ID пользователей смоделированы таким образом, что ID одной группы полностью совпадают с ID другой группы, т.е. начинаются с 1 и т.д. Поэтому будем считать, что пользователей, которые попали в две группы, нет.*
- *Считаем группу A - контрольной, а группу B - тестовой.*
- *Ожидаемые метрики, на которые должны повлиять изменения и в какую сторону, определены не были, поэтому исходим из фактически полученных результатов.*

Содержание

Итоги предобработки и проверки данных	3
Итоги анализа прохождения уровней	3
Итоги анализа получения и трат золота	5
Итоги анализа бизнес-показателей	7
Итоги проверки гипотез	9
Общий вывод и рекомендации	11

Итоги предобработки и проверки данных

В данных отсутствуют пропуски.

В большинстве колонок множество нулей, в некоторых колонках есть выбросы. В таком случае лучше сравнивать медианные значения, например, в колонках с данными о прохождении уровней, или анализировать ненулевые значения отдельно от нулевых, как в колонках с тратами и получением золота.

После удаления 1438 пользователей с ошибочными записями (16.5% от изначально количества) в группе А осталось 3568 пользователей, а в группе В - 3683. Соотношение групп пользователей составляет 49% / 51%, поэтому можно сделать вывод о том, что пользователи распределены равномерно между группами.

Итоги анализа прохождения уровней

1. В группе А пользователи максимально доходили до 24 уровня, а в группе В - до 25. Иными словами, в тестовой группе медианное значение на 4% больше относительно контрольной группы. При этом доля пользователей, не проходивших ни одного уровня, отличается лишь на 1 процентный пункт: 11.7% в группе А и 10.6% в группе В.
2. Медианное количество:
 - всех стартов в группе А составило 15, а в группе В - 17, разница показателя тестовой группы по отношению к контрольной составляет 13.3%.
 - чистых стартов в группе А равно 12, а в группе В - 14. Показатели отличаются на 16.7%. Доля пользователей, не стартовавших без бонусов в контрольной группе составляет 12.5%, а в группе В - 11.3%.
 - стартов с бонусом в обеих группах равно 2, доля тех, кто ни разу не стартовал с бонусом в контрольной группе равно 30.2%, а в тестовой группе - 27.1%.

- всех побед в группе А составило 8, а в группе В - 9, разница показателей составляет 12.5%. Доля пользователей не завершивших ни одного уровня победой в группе А составила - 16.6%, а в группе В - 14.8%.
 - чистых побед на пользователя в группе А и в группе В равно 7. При этом доля пользователей, не завершивших успешно ни один уровень, в группе А составляет 17.4%, а в группе В - 15.6%.
 - побед с бонусами одинаково в обеих группах - 1. Доля пользователей, не завершивших ни одного уровня с бонусами в группе А составила 39.5%, в группе В - 36%.
3. И в контрольной, и в тестовой группах медианное количество попыток для прохождения уровня не отличается и составляет:
- для всех уровней - 1.9;
 - для чистых уровней - 2;
 - для уровней с бонусами - 1.3.
4. Доля стартов и побед с преимуществами от общего числа стартов и побед пользователей, которые начали прохождение хотя бы одного уровня в группах не отличается, и составляет 11.1% и 14.3% соответственно.
5. Корреляционная матрица показала, что время восстановления жизни не коррелирует с каким-либо показателем прохождения уровня.

Таким образом, уменьшение времени восстановления жизни способствует небольшому увеличению количества прохождений уровней как удачных, так и неудачных.

Итоги анализа получения и трат золота

Медианное суммарное значение:

1. получения золота:

- за просмотр рекламы остался на уровне 200 в обеих группах. Процент нулевых значений в группе А составляет 87%, а в группе В - 86%.
- за прохождение глав в контрольной группе составляет 340, в тестовой - 360. Доля пользователей, не прошедших ни одну главу, в группе А составляет 76.5%, а в группе В - 73%.
- из покупки не сопоставимо, поскольку покупку золота в обеих группах совершил лишь 1% пользователей.
- за логин в Facebook не изменилось, поскольку является разовым и фиксированным. Доля пользователей, получивших золото таким образом, в группах также не отличается и составляет 9%.
- за отправку жизней в команде в группах не изменилось и составило 60. В тестовой группе жизни отправляло меньшее количество пользователей - 6.3%, чем в контрольной группе - 8.2%.
- за прохождение tutorиала для команды не отличается и равно 100 (возможно, за неделю жизни была возможность пройти только один tutorиал), как и не отличается процент тех, кто tutorиал не завершил - 71% в обеих группах.

2. трат золота:

- на покупку жизней для бонусных глав не сравнимы, поскольку в группе В никто их не приобретал, а в группе А - только один пользователь.
- на покупку бонусов в группе А составило 1400, в группе В - 900, разница в 35%. При этом не стоит утверждать, что изменения повлияли на данный показатель, т.к. эту покупку совершило 4% в каждой группе.

- на покупку бустеров в обеих группах не отличается и равно 1900, как и доля пользователей, которая совершила такую покупку - менее 1%.
- на покупку жизней в группах одинаковое - 1000. Доля пользователей, которые не приобрели жизни, в группе А составила 96.4%, а в группе В - 99.3%.
- на покупку ходов и в тестовой, и в контрольной группах составило 2590. Доля пользователей, не купивших дополнительный ход, в группе В меньше, чем в группе А - 56.5% и 59.3% соответственно.
- на создание команды в обеих группах составляет 100 (вероятнее всего, игроку не нужно создавать две и более команд, а многие присоединяются к уже существующим). Доля создавших команду также одинаково и равно менее 1%.

Корреляционная матрица показала, что время восстановления жизни не коррелирует с каким-либо показателем, связанным с оборотом золота.

В целом, уменьшение времени восстановления жизни с 30 до 3 минут не повлияло на медианные показатели, связанные с обращением золота.

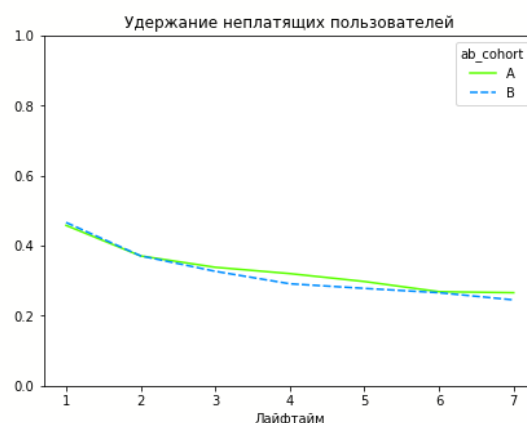
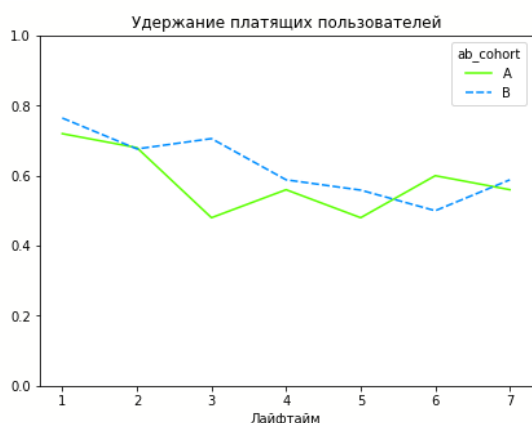
Можно отметить:

- уменьшение доли пользователей, не прошедших ни одной главы, на 3.5 процентных пункта, что связано с увеличением количества проходимых уровней, отмеченных ранее;
- уменьшение доли пользователей, отправляющих жизни в команду, на 2 процентных пункта;
- уменьшение доли пользователей, приобретающих жизни, на 3 процентных пункта;
- увеличение доли пользователей, приобретающих дополнительный ход, на 2.8 процентных пункта.

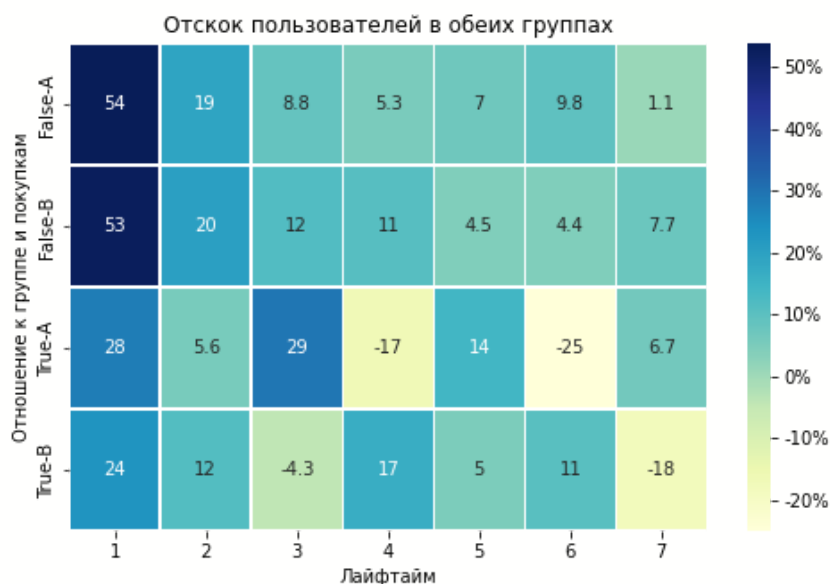
Итоги анализа бизнес-показателей

Для платящих пользователей цифры могут не отражать реальной картины, поскольку таких пользователей очень мало. К выводам для них следует относиться с осторожностью.

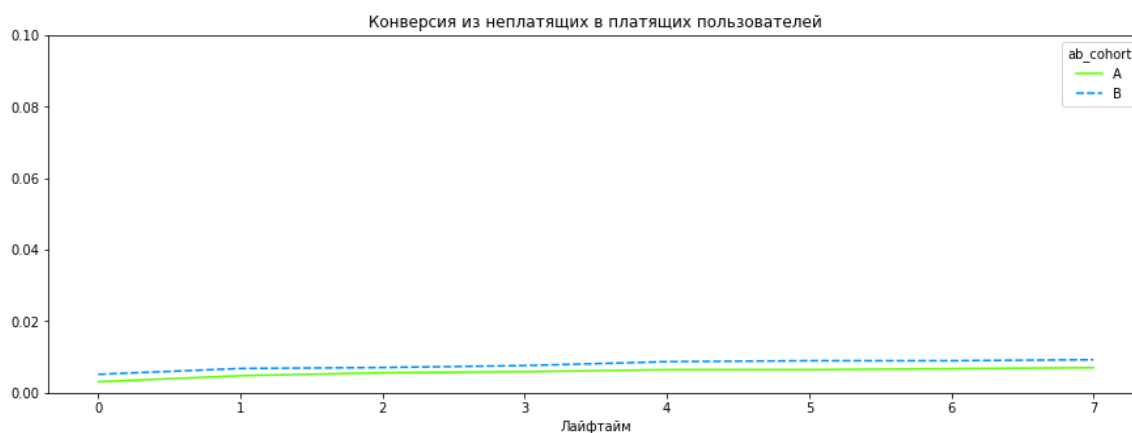
1. Удержание неплатящих пользователей в группах остается примерно на одинаковом уровне в течение всего лайфтайма. В контрольной группе на 7 день заходило 26.5% пользователей, а в тестовой - 24.5%.
2. Удержание платящих пользователей у группы В выше группы А в течение всего лайфтайма, кроме 6 дня. В группе А показатель на 7 день лайфтайма составил 56%, а в группе В - 58.8%.



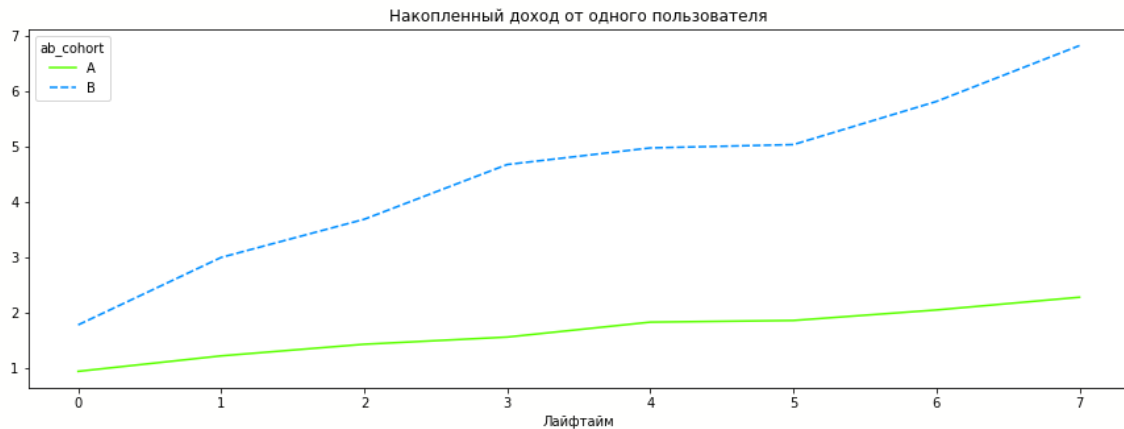
3. Отскок неплатящих пользователей в обеих группах остается примерно на одинаковом уровне в течение всего лайфтайма. В первый день после установки в приложение не зашло 53% пользователей в группе А и 54% в группе В. Процент отскока на 7 день в контрольной группе составил 73.4%, а в группе В - 75.5%.
4. Сравнить динамику отскока группы А и группы В для платящих пользователей затруднительно. Тепловые карты выглядят как противоположность друг другу. На следующий день после установки в приложение не зашло 28% пользователей в группе А и 24% в группе В. Общий процент отскока в группе А больше, чем в группе В, и равен 44% и 41% соответственно.



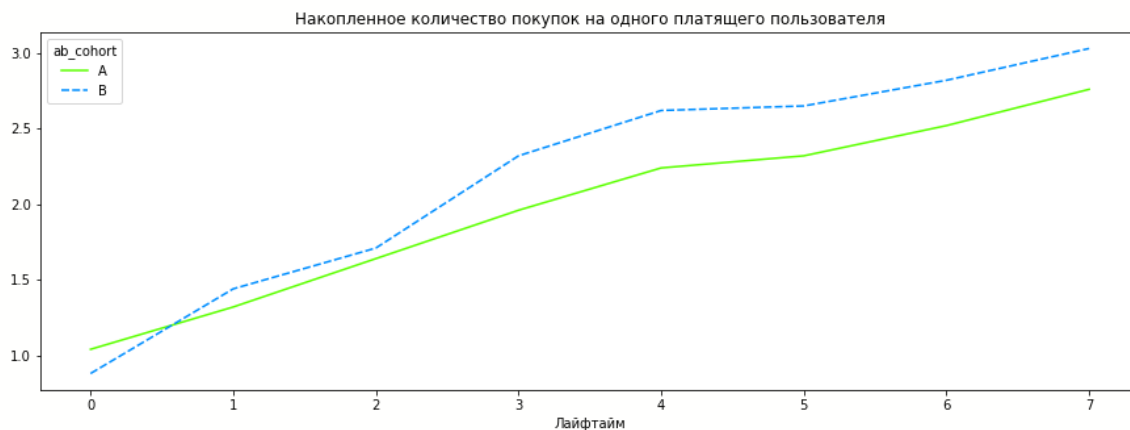
5. Конверсия из неплатящих в платящих пользователей у группы В выше, чем у группы А на протяжении всего лайфтайма. При этом за 7 дней пользования в контрольной группе доля платящих пользователей составила 0.7%, а в группе В - 0.9%.



6. Показатель LTV группы В больше группы А на протяжении всего лайфтайма. К 7 дню доход на одного пользователя в контрольной группе составил 2.28 у.е, а в группе В - 6.83 у.е. При известном САС можно было бы определить, окупались ли привлеченные пользователи.



7. Средний платящий пользователь из группы В за 7 дней лайфтайма совершил 3 покупки, а из группы А - 2.7.



Итоги проверки гипотез

H0: количество пройденных уровней пользователями в группе А и в группе В не отличается.

H1: количество пройденных уровней пользователями в группе А и в группе В отличается.

Разница в количестве пройденных уровней в группе А и в группе В является статистически значимой и составляет 4% (в группе А пользователи максимально доходили до 24 уровня, а в группе В - до 25).

H0: количество стартов уровней в группе А и в группе В не отличается.

H1: количество стартов уровней в группе А и в группе В отличается.

Разница в количестве стартов уровней в группе А и в группе В является статистически значимой и составляет 13.3% (в группе А - 15 стартов на пользователя, а в группе В - 17).

H0: полученный от пользователей доход в группе А и в группе В не отличается.

H1: полученный от пользователей доход в группе А и в группе В отличается.

Несмотря на то, что к 7 дню доход на одного пользователя в контрольной группе составил 2.28 у.е, а в группе В - 6.83 у.е., тест показал, что она не является статистически значимой.

Общий вывод и рекомендации

Прохождение уровней

После предобработки в группе А осталось 3568 пользователей, а в группе В - 3683. Соотношение групп пользователей составляет 49% / 51%, т.е. пользователи распределены равномерно между группами.

В группе А пользователи максимально доходили до 24 уровня, а в группе В - до 25. Иными словами, в тестовой группе медианное значение на 4% больше относительно контрольной группы. Тест Манна-Уитни показал, что разница статистически значима.

Доля пользователей, не проходивших ни одного уровня, отличается на 1 процентный пункт: 11.7% в группе А и 10.6% в группе В.

Таким образом, уменьшение времени восстановления жизни способствует увеличению количества прохождений уровней как удачных, так и неудачных.

Обращение золота

Было отмечено следующее:

- уменьшение доли пользователей, не прошедших ни одной главы, на 3.5 процентных пункта, что связано с увеличением количества проходимых уровней;
- уменьшение доли пользователей, отправляющих жизни в команду, на 2 процентных пункта;
- уменьшение доли пользователей, приобретающих жизни, на 3 процентных пункта;
- увеличение доли пользователей, приобретающих дополнительный ход, на 2.8 процентных пункта.

Уменьшение времени восстановления жизни с 30 до 3 минут не повлияло на медианные показатели, связанные с обращением золота, в частности, на получение золота из покупки и за регистрацию в Facebook.

Бизнес-показатели

Удержание неплатящих пользователей в группах остается примерно на одинаковом уровне в течение всего лайфтайма. В контрольной группе на 7 день заходило 26.5% пользователей, а в тестовой - 24.5%.

Отскок неплатящих пользователей в обеих группах остается примерно на одинаковом уровне в течение всего лайфтайма. В первый день после установки в приложение не зашло 53% пользователей в группе А и 54% в группе В. Процент отскока на 7 день в контрольной группе составил 73.4%, а в группе В - 75.5%.

Удержание платящих пользователей у группы В выше группы А в течение всего лайфтайма, кроме 6 дня. В группе А показатель на 7 день лайфтайма составил 56%, а в группе В - 58.8%.

На следующий день после установки в приложение не зашло 28% пользователей в группе А и 24% в группе В. Общий процент отскока в группе А больше, чем в группе В, и равен 44% и 41% соответственно.

Конверсия из неплатящих в платящих пользователей у группы В выше, чем у группы А на протяжении всего лайфтайма. При этом за 7 дней пользования в контрольной группе доля платящих пользователей составила 0.7%, а в группе В - 0.9%.

Средний платящий пользователь из группы В за 7 дней лайфтайма совершил 3 покупки, а из группы А - 2.7.

Показатель LTV группы В больше группы А на протяжении всего лайфтайма. К 7 дню доход на одного пользователя в контрольной группе составил 2.28 у.е, а в группе В - 6.83 у.е. При известном САС можно было бы определить, окупались ли привлеченные пользователи. В то же время тест Манна-Уитни показал, то разница не является статистически значимой.

Итог и рекомендации

Уменьшение времени восстановления жизни с 30 до 3 минут положительно повлияло на пользователей, но не факт, что влияние на получение прибыли такое же.

В тестовой группе отмечается увеличение количества пользователей, которые прошли хотя бы один уровень, и увеличение количества пройденных уровней в целом. При этом отскок и удержание среди неплатящих пользователей в группе В показывает результаты немного хуже, чем в группе А.

Показатели удержания и отскока платящих пользователей в тестовой группе лучше, чем в контрольной. В группе В конверсия в платящих пользователей лучше, отмечается большее количество покупок на пользователя и больший доход. Из-за того, что платящие пользователи составляют менее процента от всего количества, нельзя утверждать, что данные изменения отражают реальную картину. К тому же, несмотря на отличие в доходах от одного пользователя в 3 раза, тест Манна-Уитни показал, то разница не является статистически значимой.

Если учесть, что основной доход игры - это показ рекламы, то можно предположить, что увеличение медианного количества попыток прохождения уровней на 13.3% увеличивает доход с показа рекламы, т.к. она демонстрируется после прохождения каждого уровня.

В то же время количество полученного золота за просмотр рекламы не изменилось, как и не изменилась доля пользователей, которая этим воспользовалась.

Таким образом, если доход с показа рекламы пользователям в указанный период для данных пользователей увеличился, то A/B-тест можно считать удачным. В противном случае, таким его назвать нельзя.

Чтобы избежать подобной неоднозначности, следует заранее определить метрики и их ожидаемые изменения, по которым A/B-тест будет считаться удачным, и исходя из этого определять размер групп и набор данных для анализа. Также отдельно стоит обращать внимание на платящих

пользователей. Возможно, стоит увеличить выборку и время проведения теста, чтобы увеличить количество платящих пользователей, принимающих участие в нем. Это будет способствовать получению более точных результатов согласно закону больших чисел.