

Аналитическая записка

Содержание:

[Задание 1](#)

[Задание 2](#)

[Приложение](#)

Задание 1

Представьте, что вы аналитик в компании, которая разрабатывает приложение для обработки и оформления фотографий в формате Stories (например, для дальнейшего экспорта в Instagram Stories).

Был проведен A/B тест: тестовой группе предлагалась *новая модель оплаты коллекций шаблонов*, контрольной – *старая механика*. Ваша основная задача: проанализировать итоги эксперимента и решить, нужно ли выкатывать новую модель на остальных пользователей.

Изменения в приложении могут повлиять на такие метрики продукта как:

- Конверсия в покупку (CR),
- Средняя выручка на пользователя (ARPU),
- Средняя выручка на платящего пользователя (ARPPU),
- Количество платящих пользователей (PU),
- Количество повторных покупок,
- Количество покупок на пользователя.

По тем данным, которые приведены к A/B тесту мы можем проанализировать конверсию в покупку(CR), среднюю выручку на пользователя(ARPU), средняя выручку на платящего пользователя (ARPPU), количество платящих пользователей.

Количество повторных покупок и покупок на пользователя нельзя подсчитать, так как по данным у нас уникальные пользователи, совершившие или не совершившие покупку.

В описании A/B теста, который был проведен, нет однозначного понимания какой эксперимент проводился, какое изменение в приложении тестировалось.

Можно предположить о двух возможных изменениях:

1. Техническая деталь в оплате (например: убрали CAPTCHA или добавили кнопку)
2. Изменение модели оплаты (например: когда стала возможность покупать 10 фильтров по цене 6).

Поэтому решение об “Выкатывании” версии на всех пользователей можно оценить по двум метрикам. Если это техническое изменение в процессе оплаты, то можно рассматривать конверсию в покупку (CR) для принятия решения, а если изменение финансовой модели, то можно рассмотреть значимы изменения среднего чека на платящего пользователя (ARPPU).

Рассмотрим гипотезы о наличие различий значений метрик для тестовой и контрольной группы по всем данным и в разрезе по платформам. При наличии значимых различий мы можем говорить о том, что есть разница между тестовой и контрольной группой, и принимать решение об выкатывании новой модели на всех пользователей.

Вывод: При рассмотрении гипотез значимыми различиями оказались только в показатели ARPPU на всех данных и по отдельным платформам, все остальные различия по метрикам CR и ARPU между группами оказались не значимыми. Поэтому однозначного вывода по внедрению изменений нельзя сделать, но можно предположить, что если у нас изменилась система оплаты и она стала более

привлекательной для платящих пользователей, то можно внедрять на всех, но стоит задуматься о том, чтобы большее количество пользователей становилась платящими, так как показатели CR и ARPU статистически значимо не изменились при проведении A/B теста .

Подробнее каждая гипотеза приведена в [приложении](#).

Задание 2

Одной из основных задач аналитика является не только построение моделей, но и создание дашбордов, которые позволяют отслеживать изменения метрик и принимать на их основе оптимальные решения. Ваш руководитель хочет узнать, как обстоят дела с использованием приложения и вовлеченностью пользователей, и очень просит спроектировать ему дашборд.

Дашборд на основе данных первого задания: App performance in A/B test
https://public.tableau.com/app/profile/anna3018/viz/FinalProject_16557281528340/Dashboard1

Так как приведенные выше данные из первого задания не отвечают на вопрос изменения метрик, то были использованы придуманные (RANDOM) данные основных метрик, на основании которых можно принимать решения по вопросам вовлеченности пользователей.

Таковыми метриками считаю CR, ARPU, ARPPU

Спроектированный дашборд на данных, которые показали бы динамику изменений в CR, ARPU и ARPPU:

<https://public.tableau.com/app/profile/anna3018/viz/Illusion/Dashboard1>

Данный дашборд был спроектирован таким образом, что можно посмотреть какая ситуация на сегодняшний день в приложении с точки зрения вовлеченности пользователей, какой CR, ARPU и ARPPU. При этом можно сравнить текущие показатели со средними по месяцам и годам в прошлых периодах, а также посмотреть динамику CR в текущем месяце. График CR полезен тем, что можно посмотреть в какие дни был показатель высокий, в какие низкий, такие дни стоит рассмотреть отдельно, чтобы понять почему рос или падал данный показатель. Какие изменения привели к значительным изменениям в CR.

Приложение

По всем данным:

1. CR

CR_all, %: 0.72543

CR_A, %: 0.72824

CR_B, %: 0.72475

H0: CR в тестовой группе A не значимо отличается от CR в группе B

H1: CR в тестовой группе B значимо отличается от CR в группе B

Использовали тест ХИ2, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H0, что означает, что нет статистически значимой разницы между CR группы A и группы B, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

2. Средняя выручка на всех пользователей ARPU

ARPU_all: 7.76

ARPU_A: 6.34

ARPU_B: 8.11

H0: Средняя выручка (ARPU) на пользователя в тестовой группе A не значимо отличается от средней выручки на пользователя в группе B

H1: Средняя выручка на пользователя в тестовой группе A значимо отличается от средней выручки на пользователя в группе B

Использовали тест Манн-Уитни, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H0, что означает, что нет статистически значимой разницы между ARPU группы A и ARPU группы B, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

3. Средняя выручка на всех платящего пользователя ARPPU

ARPPU_all: 1069.75

ARPPU_A: 870.67

ARPPU_B: 1118.84

H0: Средняя выручка на платящего пользователя в тестовой группе A не значимо отличается от средней выручки на платящего пользователя в группе B

H1: Средняя выручка на платящего пользователя в тестовой группе A значимо отличается от средней выручки на платящего пользователя в группе B

Использовали Bootstrap для среднего, уровень значимости 5%.

На 5% уровне значимости мы отклоняем H0, что означает, что есть статистически значимая разница между ARPPU группы A и ARPPU группы B, значит мы принимаем решение о выкатывании новой модели на всех пользователей. Тест Манн-Уитни подтвердил статистически значимое различие между группами A и B по параметру ARPPU.

По разным платформам (iOS, Android):

4. CR_IOS

CR_all_IOS,%: 4.94624

CR_A_IOS,%: 5.53977

CR_B_IOS,%: 4.80769

H0: CR в тестовой группе A не значимо отличается от CR в группе B

H1: CR в тестовой группе B значимо отличается от CR в группе B

Использовали тест ХИ2, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H0, что означает, что нет статистически значимой разницы между CR группы A и группы B, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

5. CR_Android

CR_all_Android,%: 4.50119

CR_A_Android,%: 4.67626

CR_B_Android,%: 4.46264

H0: CR в тестовой группе A не значимо отличается от CR в группе B

H1: CR в тестовой группе B значимо отличается от CR в группе B

Использовали тест ХИ2, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H0, что означает, что нет статистически значимой разницы между CR группы A и группы B, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

6. Средняя выручка на всех пользователей ARPU_IOS

ARPU_all_IOS: 60.09

ARPU_A_IOS: 47.87

ARPU_B_IOS: 62.94

H0:Средняя выручка (ARPU) на пользователя в тестовой группе A не значимо отличается от средней выручки на пользователя в группе B

H1:Средняя выручка на пользователя в тестовой группе A значимо отличается от средней выручки на пользователя в группе B

Использовали тест Манн-Уитни, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H0, что означает, что нет статистически значимой разницы между ARPU группы A и ARPU группы B, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

7. Средняя выручка на всех пользователей ARPU_Android

ARPU_all_Android: 53.71

ARPU_A_Android: 47.84

ARPU_B_Android: 55.0

H0:Средняя выручка (ARPU) на пользователя в тестовой группе A не значимо отличается от средней выручки на пользователя в группе B

H1:Средняя выручка на пользователя в тестовой группе A значимо отличается от средней выручки на пользователя в группе B

Использовали тест Манн-Уитни, уровень значимости 5%.

На 5% уровне значимости мы не отклоняем H_0 , что означает, что нет статистически значимой разницы между ARPU группы А и ARPU группы В, значит мы не принимаем решение о выкатывании новой модели на всех пользователей.

8. Средняя выручка на всех платящего пользователя ARPPU_IOS

ARPPU_all_IOS: 1214.84

ARPPU_A_IOS: 864.1

ARPPU_B_IOS: 1309.18

H_0 :Средняя выручка на платящего пользователя в тестовой группе А не значимо отличается от средней выручки на платящего пользователя в группе В

H_1 :Средняя выручка на платящего пользователя в тестовой группе А значимо отличается от средней выручки на платящего пользователя в группе В

Использовали Bootstrap для среднего, уровень значимости 5%.

На 5% уровне значимости мы отклоняем H_0 , что означает, что есть статистически значимая разница между ARPPU группы А и ARPPU группы В, значит мы принимаем решение о выкатывании новой модели на всех пользователей. Тест Манн-Уитни подтвердил статистически значимое различие между группами А и В по параметру ARPPU.

9. Средняя выручка на всех платящего пользователя ARPPU_Android

ARPPU_all_Android: 1193.17

ARPPU_A_Android: 1023.03

ARPPU_B_Android: 1232.44

H_0 :Средняя выручка на платящего пользователя в тестовой группе А не значимо отличается от средней выручки на платящего пользователя в группе В

H_1 :Средняя выручка на платящего пользователя в тестовой группе А значимо отличается от средней выручки на платящего пользователя в группе В

Использовали Bootstrap для среднего, уровень значимости 5%.

На 5% уровне значимости мы отклоняем H_0 , что означает, что есть статистически значимая разница между ARPPU группы А и ARPPU группы В, значит мы принимаем решение о выкатывании новой модели на всех пользователей. Тест Манн-Уитни подтвердил статистически значимое различие между группами А и В по параметру ARPPU.

