

GeekBrains

Урок 5

A/B эксперименты

Зачем
нужны
эксперимен
ты?

Зачем
нужны
эксперимен
ты?

Парадокс познания



Что важно для эксперимента?

- понимание, что мы хотим узнать
- какое решение мы хотим принять
- стоимость решения
- стоимость эксперимента

С чего начинается эксперимент?

С чего начинается эксперимент?

Гипотеза

Гипотеза ЭТО...

Я верю, что ...<описание изменения>... Чтобы проверить, мы ...<описание эксперимента>... и измерим ...<описание метрики>... Эксперимент удачен, если ...<значение метрики>...

Пример гипотезы для этапа привлечения

Я верю, что наши идеи стартапов для животных найдут спрос у аудитории в фейсбуке.
Чтобы проверить это, мы на каждую из идей запустим рекламную кампанию и измерим конверсию в переход на страничку стартапа и конверсию в нажатие на кнопку «Установить».
Эксперимент удачен, если мы получим конверсию 1% в переход на страничку стартапа и 10% в нажатие на кнопку «Установить» с перехода хотя бы для одной из идей.

Пример гипотезы для этапа активации

Я верю, что если на онбординге дать возможность персонализировать содержимое приложения, то приложение будет более понятно и полезно пользователю. **Чтобы проверить это, мы** превратим онбординг в заполнение профиля пользователя и настройку приложения, как это делают лидеры социальных сетей (Twitter, Medium, Facebook) **и измерим** на органической аудитории изменение количества сессий на пользователя и RR первого дня и первого месяца. **Эксперимент удачен, если** показатели увеличатся в среднем на 20%.

Пример гипотезы для этапа удержания

Я верю, что понятный доступ к истории поиска сделает приложение более полезным для аудитории. **Чтобы проверить это, мы** сделаем историю поиска видимой на первом экране и в историю поиска добавим, в том числе, поисковые запросы, вбитые пользователем. **Мы измерим** долю аудитории, которая будет использовать историю поиска по своём возвращении (событие POI Search History). **Эксперимент удачен, если** эта доля увеличится с 8% (за 1 квартал 2018 года) до 15% на iOS.

Из чего состоит гипотеза

<описание изменения> – ваша идея

Из чего состоит гипотеза

<описание изменения> – ваша идея

<описание эксперимента> – выбранный способ проверки

Из чего состоит гипотеза

<описание изменения> – ваша идея

<описание эксперимента> – выбранный способ проверки

<описание метрики> – значимое и желанное изменение

Из чего состоит гипотеза

<описание изменения> – ваша идея

<описание эксперимента> – выбранный способ проверки

<описание метрики> – значимое и желанное изменение

<значение метрики> – критерий успеха или провала

Что является гипотезой?

1. Если мы сделаем наш сайт быстрее, то заказов станет больше
2. Если мы сделаем карточки наших товаров больше, то конверсия в покупку товаров вырастит
3. Мы посчитали Retention у пользователей, которые добавили объявление в избранное, и он выше, чем у тех, кто не добавляет. Нам нужно увеличить долю пользователей, которые добавляют объявления в избранное, и тогда вырастит Retention

Что является хорошей гипотезой?

1. Если мы сделаем наш сайт быстрее, то заказов станет больше
2. Если мы сделаем карточки наших товаров больше, то конверсия в покупку товаров вырастит
3. Мы посчитали Retention у пользователей, которые добавили объявление в избранное, и он выше, чем у тех, кто не добавляет. Нам нужно увеличить долю пользователей, которые добавляют объявления в избранное, и тогда вырастит Retention

Еще немного про гипотезы

Гипотеза должна быть формализована, опровержима и обладать критерием проверки

Способы проверки гипотез

Прототипы

Коридорные тесты

UX исследования

Расчеты на салфетках

a/b эксперименты

...

Почему важно привить культуру проверки гипотез?

В Англии часто идут дожди. Лондон один из самых сухих городов Европы, там выпадает меньше осадков, чем в Риме

Лемминги кончают жизнь самоубийством, устремляясь в морскую пучину. Они просто ошибаются и путают озеро, реки, которые они могут переплыть с морем

Когда-то люди объясняли то, что с ними происходит волей богов. Они заблуждались

Почему важно привить культуру проверки гипотез?

Мир быстро меняется и мы постоянно сталкиваемся с новым и он не похож на старый мир. Единственный способ быстро изучать новый мир - это строить гипотезы и быстро их проверять

Что такое a\b эксперимент?

Метод исследования, где контрольная группа элементов сравнивается с набором тестовых групп, в которых один или несколько показателей были изменены, для того, чтобы выяснить, какие из изменений улучшают целевой показатель

Шаблон для a/b эксперимента

1. Гипотеза (идея)
2. Что делаем? Как проверяем?
3. На каких пользователях тестируем?
4. Описание метрик: что должно поменяться и как это измерим?
5. Насколько должно поменяться? Оцениваем выборку
6. Какие решения можем принять по итогам эксперимента?

Пример эксперимента

1. Если мы добавим текст «100 % конфиденциальность – мы никогда не будем отправлять вам спам!» увеличит количество подписок, так как это вызовет больше доверия со стороны посетителей.
2. Контрольная группа пользователей: не будем добавлять ей текст
Тестовая группа: добавим текст
 1. На новых пользователях и на старых с подпиской и без подписки
 2. Количество подписок, количество подписок на пользователя, конверсия в подписку
 3. Если количество значительно улучшится, то оставим текст. Если количество значительно ухудшится или не изменится, то уберем этот текст

Эксперимент с текстом

1. «100 % конфиденциальность – мы никогда не будем отправлять вам спам!» .
Количество подписок уменьшилось на 18.7%
2. «Мы гарантируем 100 % конфиденциальность. Ваши данные не будут разглашаться». Количество подписок увеличилось на 19.47%

Ухудшающие эксперименты

1. Нужен для предотвращения или подтверждения дорогой разработки
2. Эксперимент Financial times со скоростью загрузки страницы

Пример эксперимента

1. Если мы увеличим кол-во разрешений на пуши получаемых от пользователей, то вырастит их Retention
2. Контрольная группа пользователей: ничего не меняем
Тестовая группа: не запрашиваем разрешение
 1. На новых пользователях
 2. Retention
3. Если доля пользователей, которые включают разрешения влияет на Retention, то нужно сосредоточиться на увеличение доли

A/B тесты в Я.Недвижимости

1. Пользователи разбивались на бакеты в разных измерениях
2. Админка для выкатки экспериментов
3. Вручную следили за тем, чтобы на пользователях не запускались эксперименты одновременно
4. В логах и в коде страницы всегда видно запущен ли эксперимент на конкретном пользователе

Side by Side

- › Один человек видит два варианта на одной странице и выбирает тот, который больше нравится
- › Подходит для тестирования иконок, логотипов, текстов
- › Не стоит проверять функциональность

Сервис Я.Толока <https://telegra.ph/Testirovanie-dizajna-na-YAToloka-02-10>

A/A тесты

A/B тест сравниваем две разные версии на контрольной и тестовой группе пользователей

A/A – то же самое, только сравниваем одну версию. Цель – не увидеть различий в их показателях.

Про что важно помнить

Размер аудитории

Про сезонность

Про эффект новизны

Про пересечение экспериментов

Количество тестов

Сложность поддержки

Размер выборки и время проведения

Мы хотим узнать, сколько потребуется пользователей в тестовой и контрольной группе, чтобы заметить определенное изменение в метрике.


Наша текущая конверсия 20%, мы планируем ее увеличить на 2 пп, вероятность 80%


Калькулятор:

<http://www.evanmiller.org/ab-testing/sample-size.html>

GeekBrains

Question: How many subjects are needed for an A/B test?

Baseline conversion rate: %  20% [\[link \]](#)

Minimum Detectable Effect: %  18% – 22%

The Minimum Detectable Effect is the smallest effect that will be detected (1-β)% of the time.

Absolute
 Relative

Conversion rates in the gray area will not be distinguishable from the baseline.

Sample size:
6,347
per variation

Statistical power 1-β: 80% Percent of the time the minimum effect size will be detected, assuming it exists

Significance level α: 5% Percent of the time a difference will be detected, assuming one does NOT exist

See also: [How Not To Run an A/B Test](#)

Количество тестов

- › Один в рамках страницы
- › Один в рамках сценария
- › Один в рамках целевой группы
- › При небольшой аудитории лучше тестировать гипотезы последовательно

Разметка событий

Разметили новое, забыли старое

Разметили по-разному

Не протестировали разметку перед выкладкой эксперимента

Проведение эксперимента

Длительность не меньше недели

Не подглядывать в результаты раньше времени (проверка в-первый день)

Один и тот же человек видит оба варианта

Выборки не однородны

Подведение итогов

Не забываем считать доверительный интервал

Ориентируемся на заданные заранее пороговые значения

Не забывайте проверить, не ухудшил ли ваш эксперимент ключевые метрики

Важно понимать результаты, почему конверсия выросла или упала

Чек-лист

- › Цель (на какой KPI влияет)
- › Метрика и критерий успешности (как будем интерпретировать данные и принимать решение)
- › Определение размера выборки и длины эксперимента
- › Разметка событий и выборок
- › Проверка разметки и выборок
- › Проведение эксперимента
- › Сбор данных и выводы
- › Проверка успешности изменения спустя время

Чек-лист для ab эксперимента

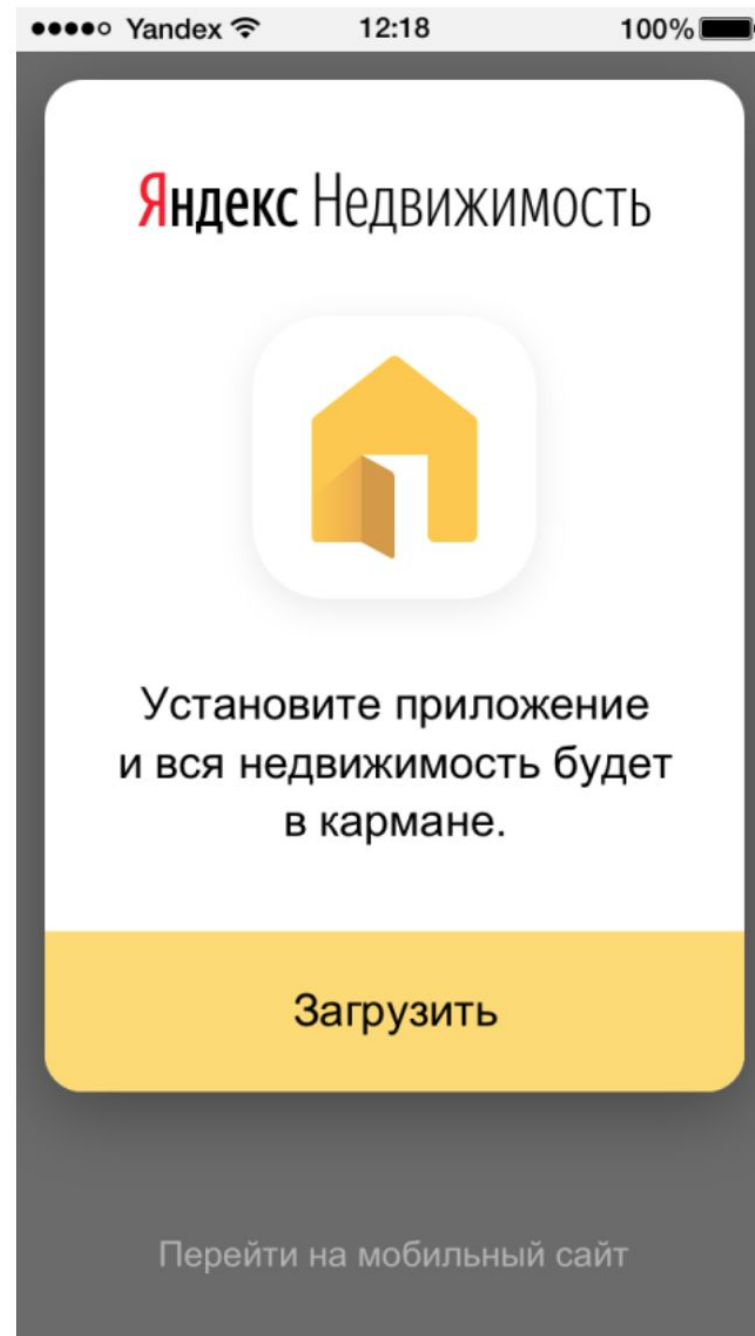
- › Цель (на какой KPI влияет)
- › Метрика и критерий успешности (как будем интерпретировать данные и принимать решение)
- › Определение размера выборки и длины эксперимента
- › Разметка событий и выборок
- › Проверка разметки и выборок
- › Проведение эксперимента
- › Сбор данных и выводы
- › Проверка успешности изменения спустя время
- › Стоимость эксперимента и стоимость решения

Кейс про Я.Недвижимость

«У Я.Недвижимости есть мобильный сайт и мобильное приложение. Конверсия в целевое действие «звонок» в мобильном приложении выше, чем на сайте.

В мобильном сайте есть баннер про мобильное приложение. Давайте его улучшим, чтобы еще больше пользователи устанавливало приложение».

Кейс про Я.Недвижимость



Стоимость эксперимента

- › Макеты – 2ч
 - › Фронтенд – 3ч
 - › Тестирование – 2ч
 - › Подвести результаты по эксперименту и требования – 1ч
- Итого: 8 часов ~ 1 рабочий день

Стоимость решения

Баннер показываем пользователю один раз при первом визите
Аудитория мобильной версии сайта – 2000 человек
Из них новых пользователей – 1000 человек

- › Конверсия в звонок в мобильной версии сайта – 5%
- › Конверсия в звонок в мобильном приложении – 6%

- › Конверсия в установку приложения с баннера – 10%
- › Конверсия в установку приложения с нового баннера –???

Стоимость решения

Баннер показываем пользователю один раз при первом визите
Аудитория мобильной версии сайта – 2000 человек
Из них новых пользователей – 1000 человек

- › Конверсия в звонок в мобильной версии сайта – 5%
- › Конверсия в звонок в мобильном приложении – 6%

- › Конверсия в установку приложения с баннера – 10%
- › Конверсия в установку приложения с нового баннера – **15-25%**

Стоимость решения

Со старым баннером, конверсия 10%

$$1000 * 0,05 + (1000 * 0,1) * 0,06 = 50 + 6 = 56 \text{ звонков}$$

С новым баннером, если конверсия 15%

$$1000 * 0,05 + (1000 * 0,15) * 0,06 = 59 \text{ звонков}$$

С новым баннером, если конверсия 25% и конверсия в звонок в таче скорее всего упадет и станет 4,5%

$$1000 * 0,045 + (1000 * 0,25) * 0,06 = 60 \text{ звонков}$$

Сколько денег потеряет Amazon, если у него упадет конверсия в покупку?

Визитов в месяц 2,6 млрд в месяц
Конверсия в покупку 10%
260 млн визитов заканчивается покупкой
1 покупка приносит 2\$

Конверсия снижается до 9%

Сколько денег потеряет Amazon, если у него упадет конверсия в покупку?

Визитов в месяц 2,6 млрд в месяц
Конверсия в покупку 10%
260 млн визитов заканчивается покупкой
1 покупка приносит 2\$

Конверсия снижается до 9%

Amazon потеряет 52 млн \$

Резюме

Для хорошего a/b эксперимента нужна хорошая гипотеза

a/b эксперименты сложный инструмент, важно использовать не только его при проверке гипотез

нужно вкладывать время в правильное проектирование эксперимента

при подведении итогов важно разобраться в том, что повлияло на результат и как повлиял эксперимент на ключевые метрики

Домашнее задание

1. выберите место, которое вы хотите улучшить в своем продукте и придумайте 5 гипотез для экспериментов
2. опишите подробно один из экспериментов, используя чек-лист, о котором было рассказано в лекции
3. посчитайте сколько будет стоить этот эксперимент и какой профит вы получите, если эксперимент будет успешным