Вариации А/Втестирования и его приложения

Цель модуля

Научиться подбирать метрики под ваш продукт, а также выбирать тип A/B-тестов.

Задачи модуля

- ✓ Познакомитесь с пирамидой метрик и факторами, влияющими на выбор метрик
- Узнаете, какие ошибки приводят к неверным А/В-тестам
- ✓ Поймёте, какие дизайны А/В-тестов существуют и как их использовать

Выбор метрики и определение цели бизнеса

Цель урока

Мы уже успели немного познакомиться с метриками, давайте погрузимся в них ещё глубже.

Задачи урока

- Рассмотрим иерархию метрик пирамиду метрик
- Расширим список основных метрик
- Разберём, как придумать метрику, которая подходит для цели А/В-тестирования

Вспомним первичные и вторичные метрики

Мы уже говорили, что не стоит забывать включать в список рассматриваемых метрик ещё и вторичные метрики. Это позволяет не только наблюдать улучшения, но и контролировать нежелательные изменения.

Вспомним первичные и вторичные метрики

Какую гипотезу вы бы посчитали успешной?

Без использования вторичной метрики можно было бы ошибиться!

Тест	Прибыль сервиса	Отток клиентов
Гипотеза # 1	+ 5,9 %	+ 0,1 %
Гипотеза # 2	+ 7,4 %	- 4,2 %

Но как вообще сформировать набор метрик?

Понять, какие метрики когда выбрать и как потом использовать полученные показатели, непросто, но за несколько экспериментов можно. Сейчас мы попробуем разобрать с вами примерные алгоритмы действий.



От чего зависит выбор метрики? От продукта!

В первую очередь при выборе метрик стоит ориентироваться на три первостепенных фактора:

- Этап развития
- Тип продукта
- Цели бизнеса

От чего зависит выбор метрики? От продукта!

Разберём каждый случай подробнее на дальнейших примерах.

Но перед тем как продолжить, давайте представим, что вы главный продукт-менеджер онлайн-школы. Какие 5 метрик более всего будут отражать успех школы?

Не спешите с ответом. Будем разбираться по ходу дела!

1/3 — метрики и этап развития

На каждом этапе развития бизнеса ему важны соответствующие метрики.



1/3 — метрики и этап развития

На каждом этапе развития бизнеса ему важны соответствующие метрики.



На каком этапе развития? Какие метрики подойдут для этого этапа?

Выбор метрик зависит не только от этапа жизненного цикла, но и от самого продукта. Например, метрики приложения, облачного сервиса и интернет-магазина тоже будут различаться, потому что у этих продуктов разные цели и задачи.

Можно условно привести четыре показателя вовлечённости, за которыми обычно следят SaaS-компании*:

- среднее время перехода пользователя с пробной на платную версию программы
- процент пользователей, которые использовали ту или иную функцию продукта
- среднее количество ключевых действий, выполненных пользователем
- среднее число ключевых действий, предпринятых за сессию

Но если за этими метриками следят в одной SaaS-компании, то это не значит, что их же будут использовать и в другой.

^{*} SaaS — это модель предоставления лицензии на программное обеспечение по подписке

Даже одна и та же метрика будет иметь разное значение для разных продуктов!

Например, есть NSM — North Star Metric, или «метрика Полярной звезды».

Она лучше всего определяет основную ценность продукта для клиентов. NSM будет разной в зависимости от продукта. Например, для Airbnb это количество бронирований, а для Facebook — показатели активности пользователей.

Выбор метрик зависит не только от этапа жизненного цикла, но и от самого продукта. Например, метрики приложения, облачного сервиса и интернет-магазина тоже будут различаться, потому что у этих продуктов разные цели и задачи.

Какие показатели вовлечённости подойдут? Возможные варианты:

- среднее время перехода пользователя от бесплатных версий курса до платных программ
- процент пользователей, которые купили тот или иной курс
- среднее количество ключевых действий, выполненных пользователем
- среднее число ключевых действий, предпринятых за сессию
- средний процент прохождения курса

У каждого бизнеса свои цели, а метрики зависят от них.

Например, одни хотят знать, сколько человек зашли на сайт почитать информацию о продукте, а другим важнее, сколько процентов пользователей из общего количества зашедших на сайт скачали приложение.

В итоге каждая команда разработки индивидуально ставит цели и задачи, выбирает, какие метрики и как часто отслеживать.





Пусть две разные команды запускают два похожих приложения. Значит ли это, что у них одни и те же цели и им подойдут аналогичные показатели?

Какие возможные цели?

- Количество людей, поступивших на курс?
- Или количество людей, прошедших курс?
- А может, количество просмотров главного сайта?

Подумайте над этим.

При этом при всём, определяясь со списком метрик, мы ещё помним, что у нас есть роли самих метрик: первичная и вторичная.

Но! Первичные и вторичные — это не всё. Есть другие роли!

При возможности можно выделить больше групп метрик, если данные вашего сервиса позволяют их посчитать.

Ключевые метрики — метрики, отвечающие за главные задачи сервиса. Их увеличение точно улучшает/ухудшает сервис.

Ограничительные метрики (вторичные) — метрики, которые нельзя «сломать».

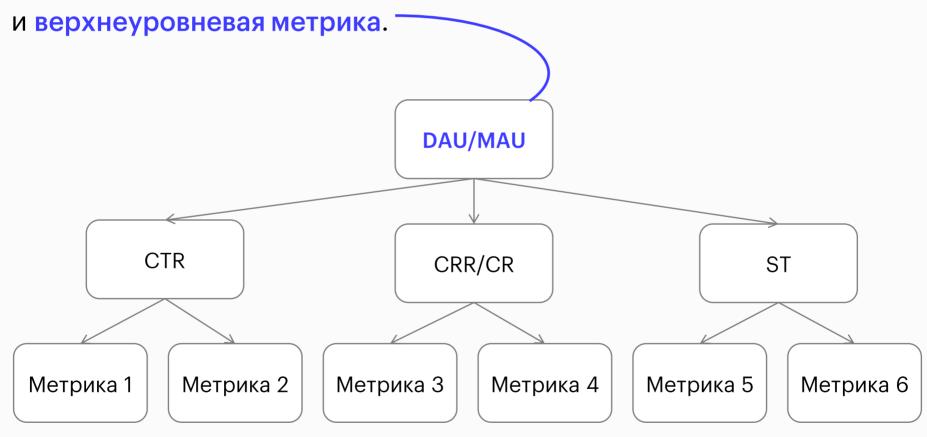
Основные критерии оценки изменений — метрики, согласованные с ключевыми метриками, но более чувствительные. Например, метрика «прибыль» очень не чувствительна к изменениям, поэтому полезно следить за более чувствительными метриками.

Главные метрики эксперимента позволяют контролировать ваш текущий эксперимент.

Хорошо, мы примерно определились со списком метрик, который будем использовать. Теперь надо их как-то приоритизировать, то есть собрать из них некоторую структуру.

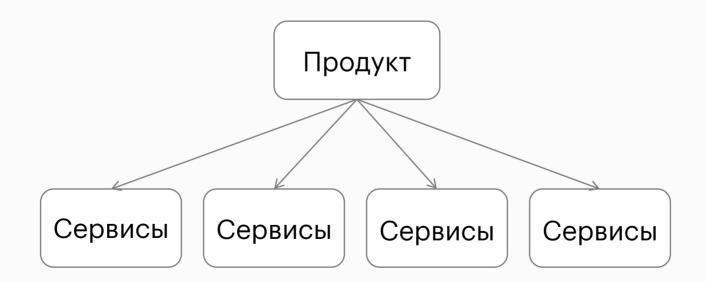
Пирамида метрик

Пирамида метрик — это древовидная структура, во главе которой находится самая основная



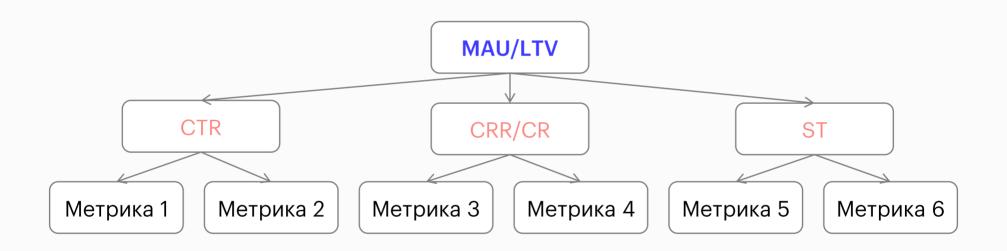
Пирамида метрик

Основная метрика отражает успешность продукта в целом, но не может нам сказать, насколько успешен сервис внутри **продукта**. Эту информацию нам смогут дать метрики конкретного сервиса.



Пирамида метрик

Составим возможную пирамиду метрик



MAU (Monthly Active Users — количество активных пользователей в месяц).

LTV (Lifetime Value — сколько каждый пользователь приносит денег).

СТР — отношение просмотров/действий.

CRR/CR — информация про возвращающихся/ уходящих пользователей.

ST — время сессии.

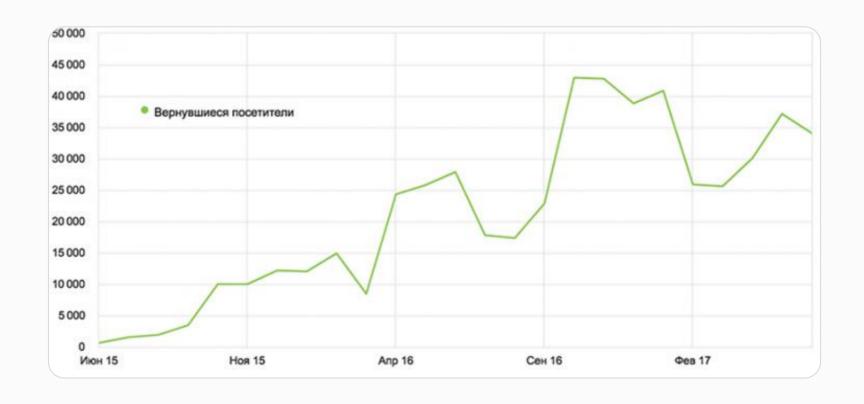
Интерпретация метрик

- Как уже говорилось, смотреть на метрики необходимо в динамике
- Смотреть на метрики в широкой перспективе: неделя/месяц/год
- Продукт может быть достаточно сложно устроен, поэтому следить за продуктом надо в целом, то есть по нескольким метрикам сразу

Интерпретация метрик

Пытаться объяснить рост или падение метрик в какой-то период — залог хорошего понимания продукта

Значимы ли полученные результаты? Логично ли изменение или есть баг?



Практическая значимость успешных тестов: кейсы

Разница в тестировании статистически значима. Этого достаточно?

Маркетинг. Провели акцию, конверсия в покупку выросла. Но акция потребовала больших финансовых затрат



- Приложение. Метрики количества активных пользователей не изменились, но опрос пользователей показал, что стало им стало комфортнее пользоваться
- Фармацевтика. От лекарства выздоравливает больше пациентов. Но есть побочные эффекты



План выбора метрики

```
Данные логируются в виде events/logs?
       Да
       Есть user id?
              Да
              Гипотеза: общее значение метрики
              Гипотеза: поведение среднего пользователя
              Нет
              Можно ли посчитать без user_id?
                     Да
                     Гипотеза: общее значение метрики
                     Нет
                     (\Xi)
       Нет
```

Тестирование не получится провести Надо добавить логирование

Итоги и выводы урока

- Метрика зависит от типа продукта, этапа развития и целей компании
- Если после экспериментов метрика
 не изменилась, то можно оставить фичу,
 если она потенциально имеет долгосрочный
 эффект или просто добавляет пользователям
 комфорта
- Следите за метриками в динамике!
- Выбрать нужную метрику сразу непросто, но за несколько экспериментов это сделать можно