

Прикладная статистика и анализ данных

Съезд 7

АВ-тестирование

Горные велосипеды. Распродажа 60%

В наличии более 1 000 моделей!

bike.ru

Распродажа! Sale! Rebajas! Saldi!

Не пропустите! До 1 числа продаем горные велосипеды по смешным ценам! :)

bike.ru



Офисные кресла

Скидка от 10 кресел - 15%! Скидка в шоу-рум до 50%! Доставка – бесплатно!!!



Офисные кресла

Скидка от 10 кресел - 15%! Скидка в шоу-рум до 50%! Доставка – бесплатно!!!

Ванны. Огромный выбор!

[Чугунные](#) [Акриловые](#) [Стальные](#)

От ведущих производителей. Быстрая доставка! Низкие цены!
Звоните!

bath.ru

Ванны. Огромный выбор!

[150x70](#) [160x70](#) [170x70](#)

От ведущих производителей. Быстрая доставка! Низкие цены!
Звоните!

bath.ru

Ванны. Огромный выбор!

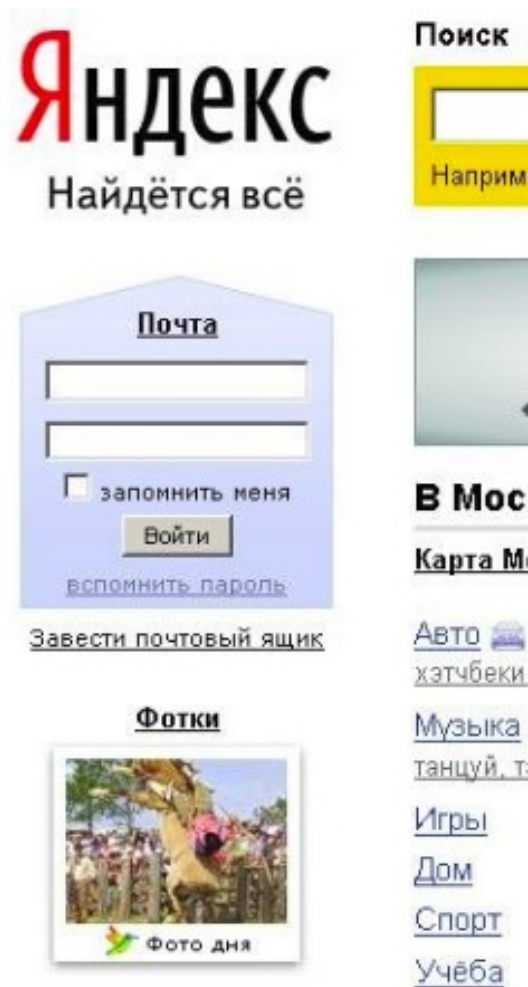
[BLB](#) [Family](#) [Jacob Delafon](#)

От ведущих производителей. Быстрая доставка! Низкие цены!
Звоните!

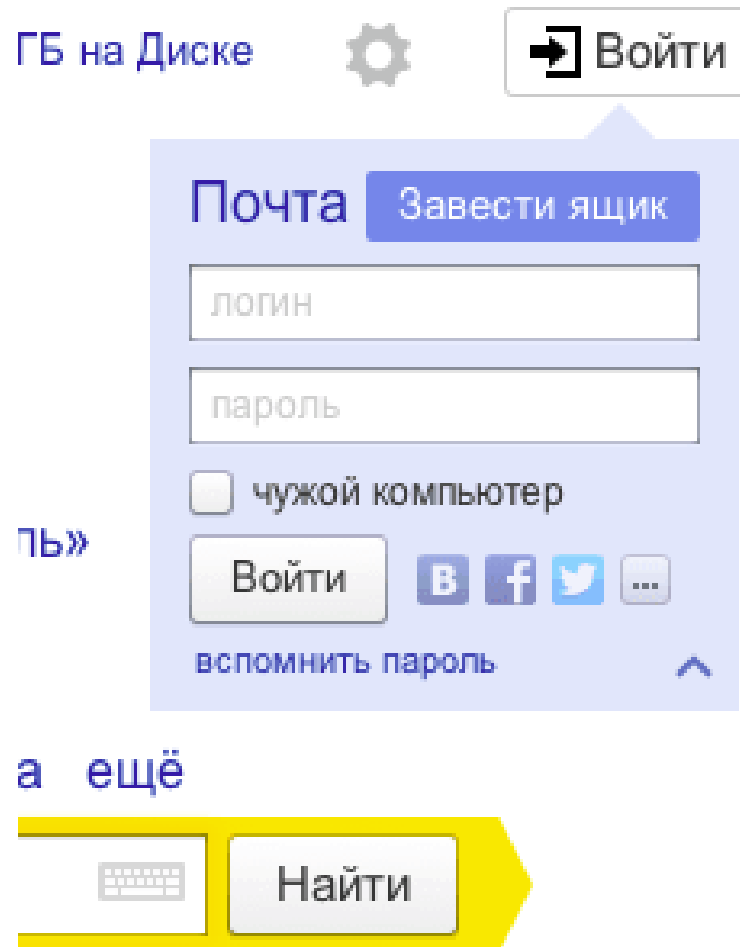
bath.ru

Дизайн входа в почту

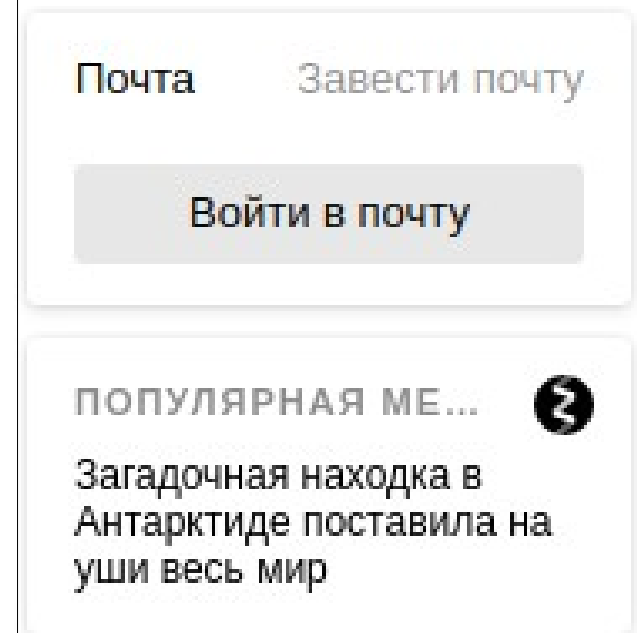
2007-2012



2013-2015

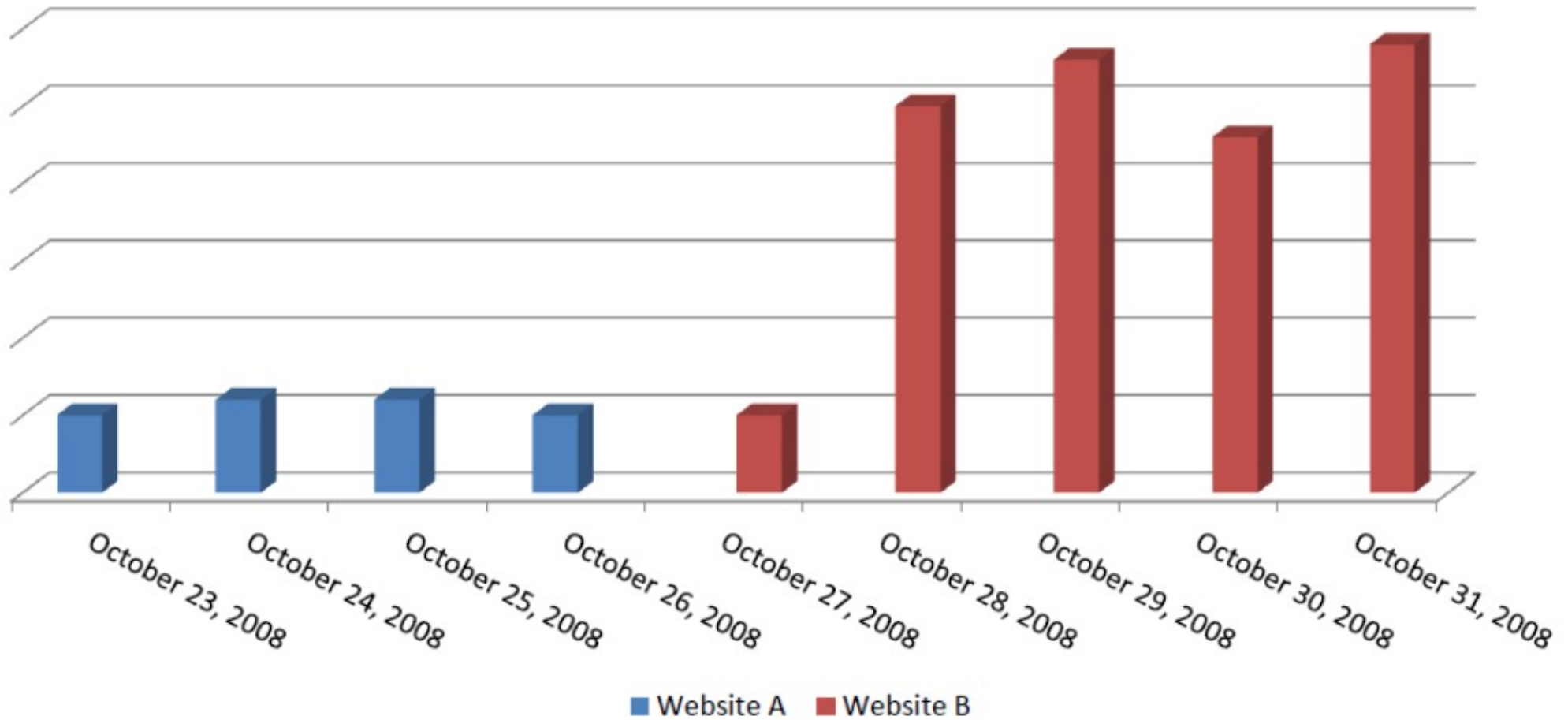


2018



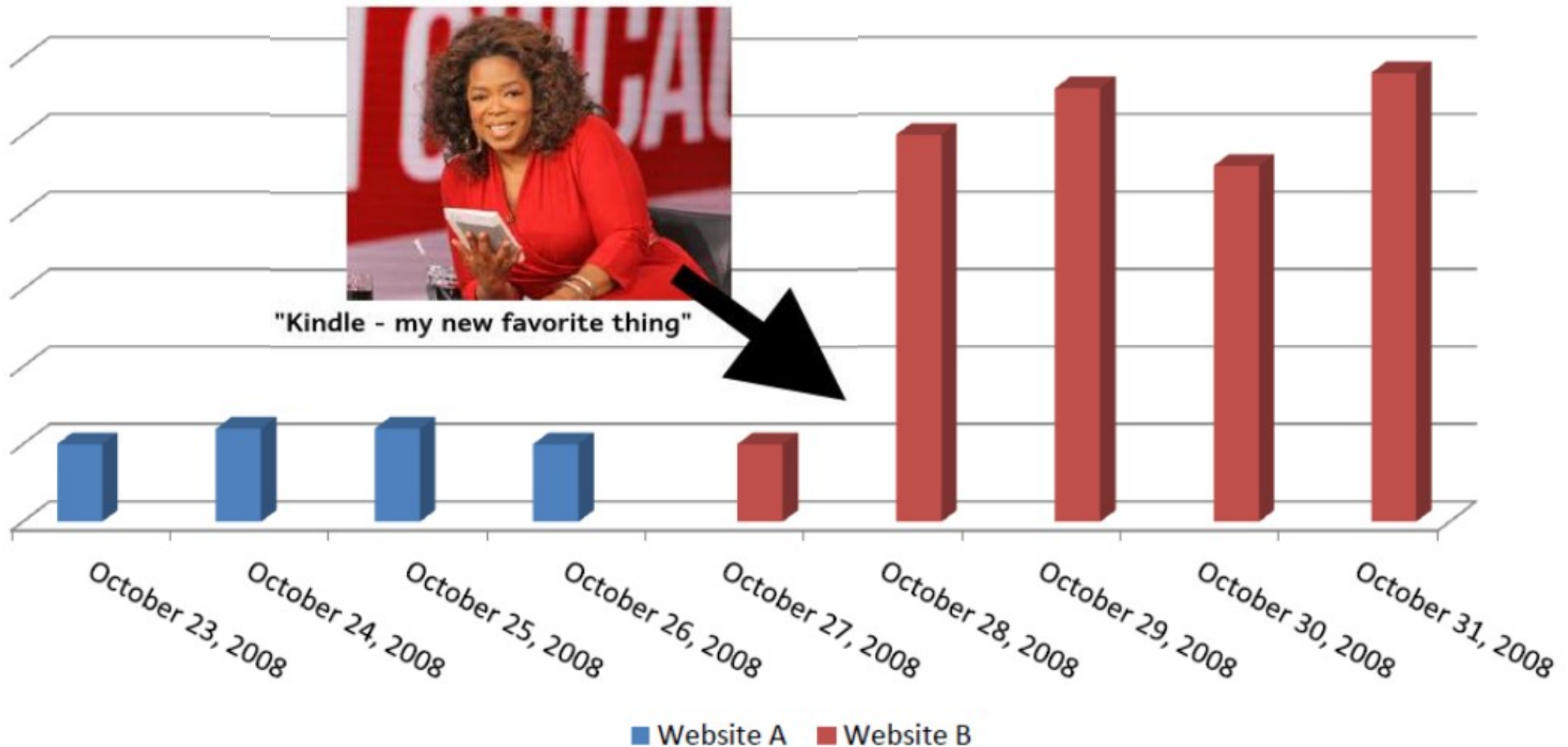
Продажи

Amazon Kindle Sales

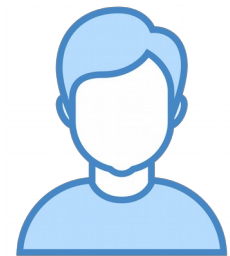
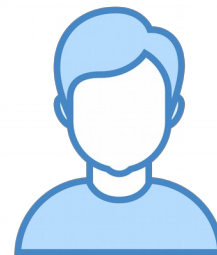
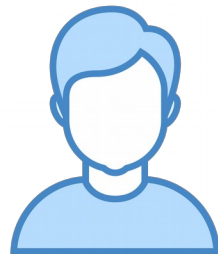
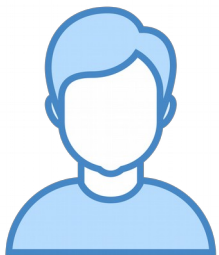
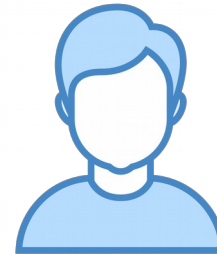
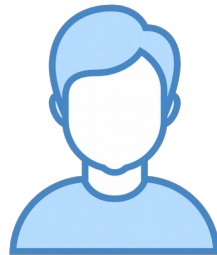
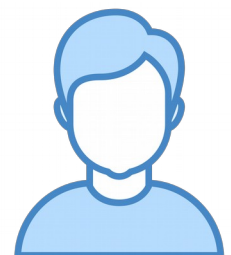
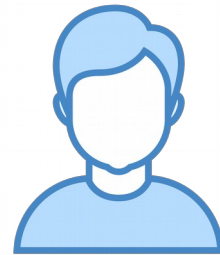
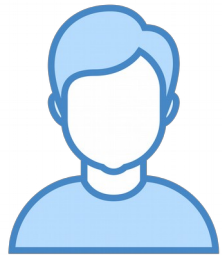


Продажи

Amazon Kindle Sales



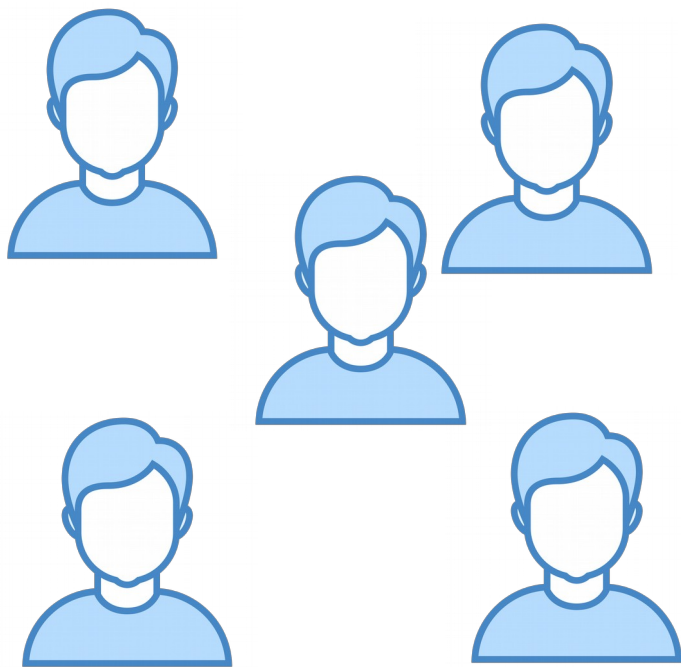
Пользователи



Группа А

Контрольный сегмент

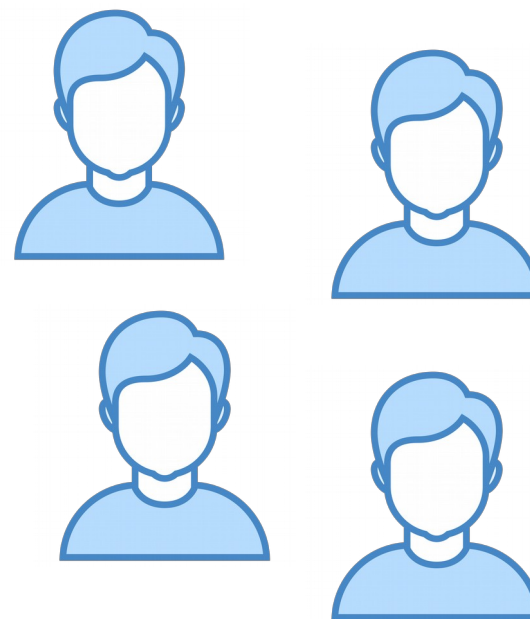
Пользователи видят
прежнюю версию сервиса



Группа В

Тестовый сегмент

Пользователи видят
измененную версию сервиса



Варианты разбиения

- По пользователям (куки)
- По визитам (сессии)
- По действиям (запросы)

В какой пропорциональности делить?

Тестовый сегмент обычно выбирается маленьким, т.к. есть риск негативного эффекта от изменений.

Типы экспериментов

1. Явные — интерфейсы, функциональность
2. Неявные — ранжирование, персонализация

Поведение пользователей:

1. Сезонность
 - Использовать периоды, кратные неделе
 - Исключать крупные праздники
2. Спецэффекты
 - Изменение множества пользователей
 - Изменение поведения пользователей
3. Поломанные пользовательские привычки

АА-тесты: проверка, что при построении функциональности нигде не налажали

Метрики

1. Конверсия = доля от общего числа пользователей, совершивших некоторое действие

- Клик: $CTR = (\text{число кликов}) / (\text{число показов})$
- Заполнение формы, регистрация
- Совершение покупки

2. Экономические метрики

- Общая прибыль
- Средний чек

3. Поведенческие факторы

- Глубина просмотра страниц
- Средняя продолжительность сессий

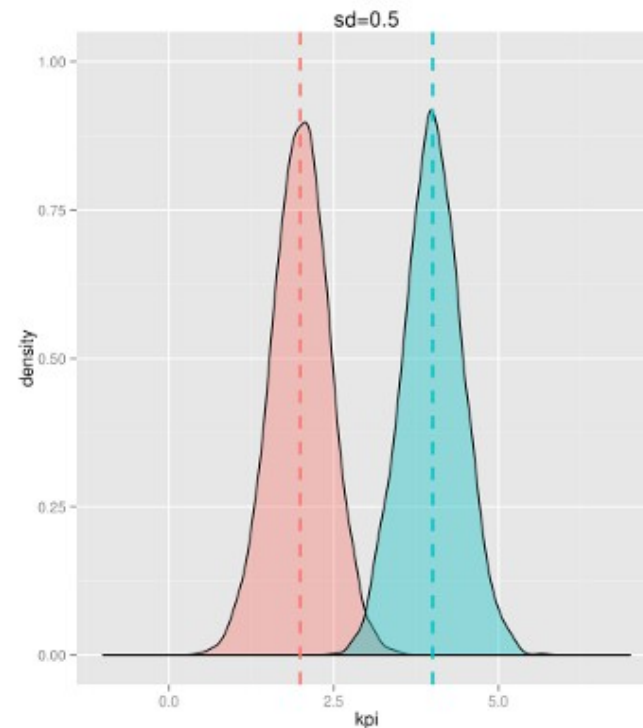
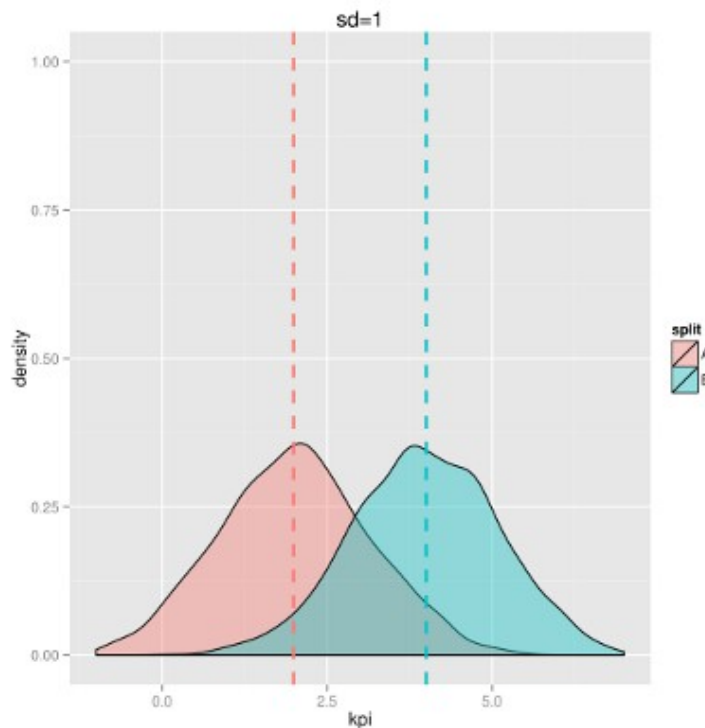
Анализ результатов

X_1, \dots, X_n — действия пользователей группы A

Y_1, \dots, Y_m — действия пользователей группы B

$m(X)$, $m(Y)$ — значения метрик

$H_0: m(X) = m(Y)$ vs. $H_1: m(X) \neq m(Y)$



Анализ результатов

Распределение	Критерии	Примеры
Нормальное	Критерий Стьюдента, критерий Аспина-Уэлша	Средняя прибыль по пользователям
Бернуллиевское	Z-критерии пред. съезда, Точный тест Фишера, Критерий хи-квадрат	CTR
Пуассоновское	Е-тест, С-тест	Количество транзакций на пользователя
Мультиномиальное	Критерий хи-квадрат	Количество приобретенных пользователем продуктов
Другое	Критерий Уилкоксона-Манна-Уитни, семплирование	

Результаты

1. Яндекс, 2017:

- почти 5000 экспериментов в год
- ~400 одновременно
- в 28% случаев изменения принимаются

2. Бинг, 201?:

- в ~30% случаев изменения принимаются

3. Гугл, 201?:

- в ~10% случаев изменения принимаются

4. В среднем ~20% случаев изменения принимаются

Сервисы

- Яндекс.Директ
- Яндекс.AppMetrica
- Google Analytics
- Vwo.com
- Список сервисов <http://boosta.ru/services/ab-testing/>
- Куча онлайн-калькуляторов

Are your results significant?

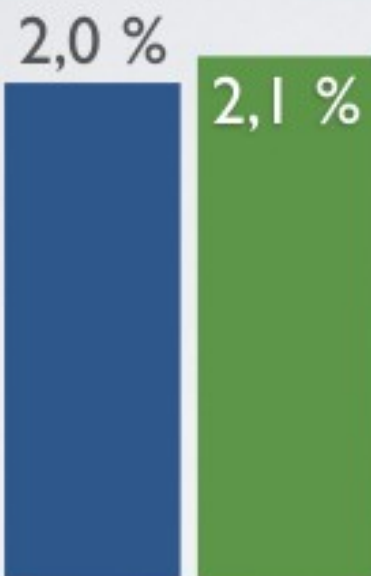
	Control	Variation
Number of Visitors	842	527
Number of Conversions	58	45

CALCULATE SIGNIFICANCE

P-Value (Range from 0-1)	Significance
0.135	No

Пример 3: сервис бронирования отелей Островок.ру

выигрыш на 5%



Конверсия

потеря на 10%



Средний чек

ТЕСТ РАЗНЫХ АЛГОРИТМОВ
СОРТИРОВКИ

Пример из Википедии

- Клиентская база 2000 человек
- Решено провести рассылку писем с промокодами
- Клиенты поделены на две равные группы
- Содержание писем одинаково, кроме фраз
 - "Offer ends this Saturday! Use code A1"
 - "Offer ends soon! Use code B1"
- Результат: в первой группе промокодом воспользовались 50 человек, во второй 30 человек.
- Результат статистически значим.

<https://vwo.com/ab-split-test-significance-calculator/>

BCE!