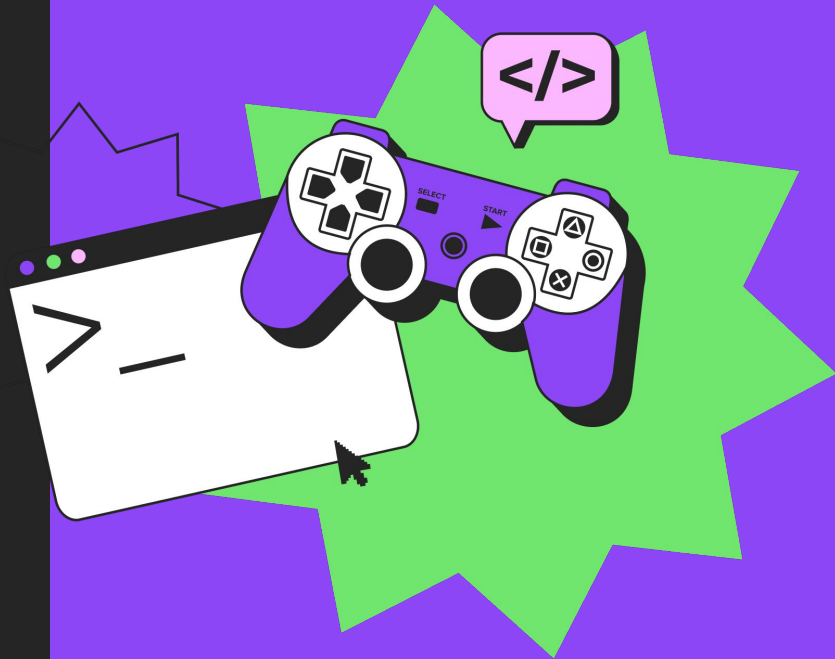


# A/B тестирование

Урок 2

## A/B тесты





# Введение в А/В тестирование



# На этом уроке мы разберем



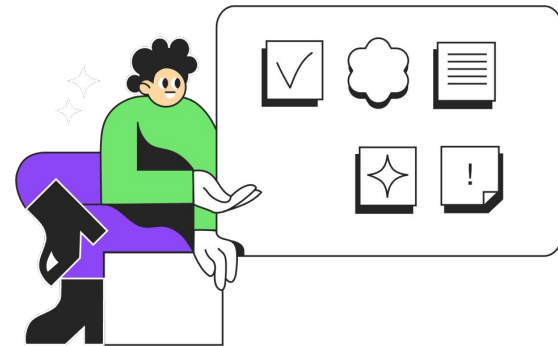
Рассмотрим какие виды экспериментов существуют



Обсудим как понять можно ли и нужно ли вам проводить А/Б-тест



Посмотрим как выглядит процесс А/В тестирования





- Введение в проблематику оценки эффекта и виды А/В тестов

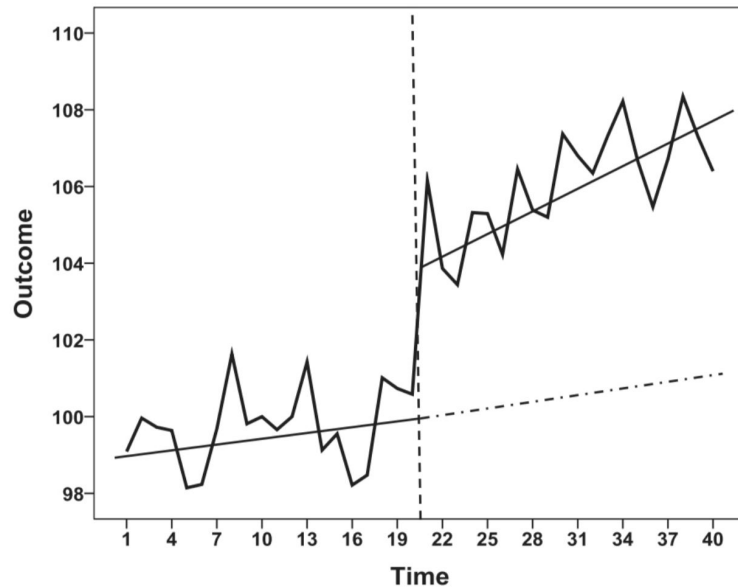


## Введение в проблематику оценки эффекта

Вы запускаете (маркетинговую кампанию /продуктовую фичу/изменение бизнес-процесса) и хотите понять эффект от этого .  
Что еще может повлиять на изменение вашей метрики кроме вашей активности?

Сезонность , тренд , продуктовая активность ваших коллег , законодательные изменения, политическая и экономическая ситуация, а также множество других факторов

Таким образом может быть сложно оценить в лоб инкрементальный эффект от изменения и вы можете ошибочно интерпретировать результаты. Ошибаться с направленностью эффекта и его размером



# A/B-эксперименты

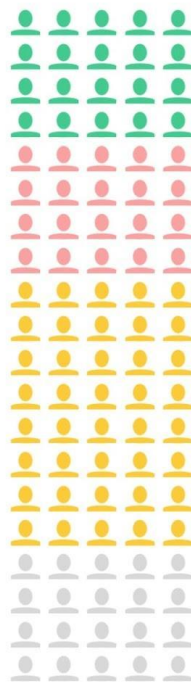
**A/B-тестирование** - это метод исследования, который используется для оценки влияния продуктовых изменений на метрики. Один из самых популярных и корректных методов для проверки гипотез.

1. Юзеров рандомизированно делят на несколько групп

Группа А(контрольная) видит текущую версию продукта  
Группа В(тестовая) видит новую версию продукта

2. Запускают эксперимент, проводя его параллельно для контрольной и тестовой групп

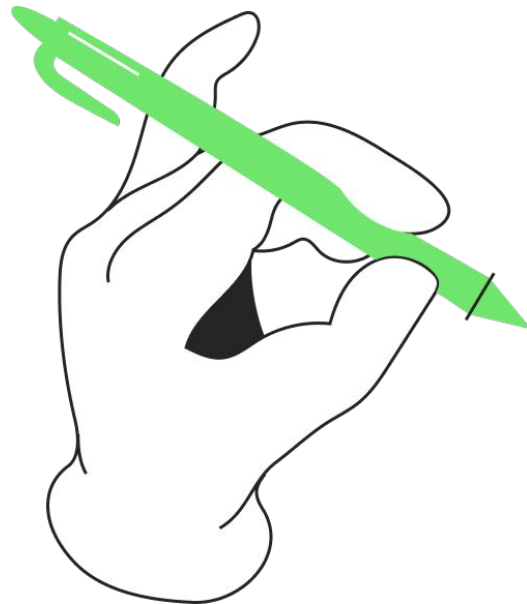
3. Сравнивают метрики в двух группах и принимают продуктовые решения



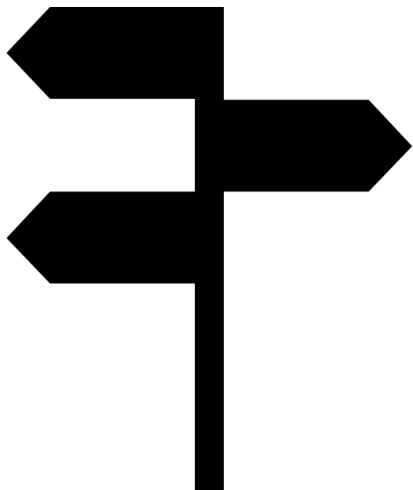


## Популярные в индустрии виды экспериментов(A/B)

- 1) A/B-тестирование (или Сплит-тестирование).
- 2) A/A-тестирование.
- 3) MVT (мультивариативное тестирование).
- 4) Ухудшающие A/B-тесты.



# А/В-эксперименты



## Направления для применения А/В тестов:

- Тестирование новых фич в приложениях и нового функционала на сайтах
- Эксперименты в оперейшнс
- Эксперименты с дизайном
- Тестирование алгоритмов
- Эксперименты с ценообразованием





## A/B-эксперименты

Плюсы:

- Понятный аппарат подсчета
- Возможность точно измерить количественный эффект
- Возможность нейтрализации внешних эффектов
- Переносимость выводов с выборки на ген совокупность

	Random assignment	No random assignment	
Random sampling	Causal and generalizable	Not causal, but generalizable	Generalizable
No random sampling	Causal, but not generalizable	Neither causal nor generalizable	Not generalizable
	Causal	Not causal	

## Когда невозможно/проблемно применить A/B тест?

- Маленький трафик
- Дорогая(долгая разработка) разработка
- Нет возможности контролировать попадание человека в контроль/тест
- Сетевые эффекты (агрегаторы такси/соцсети)
- Этические/репутационные/законодательные ограничения
- Нет инфраструктуры/компетенций в команде

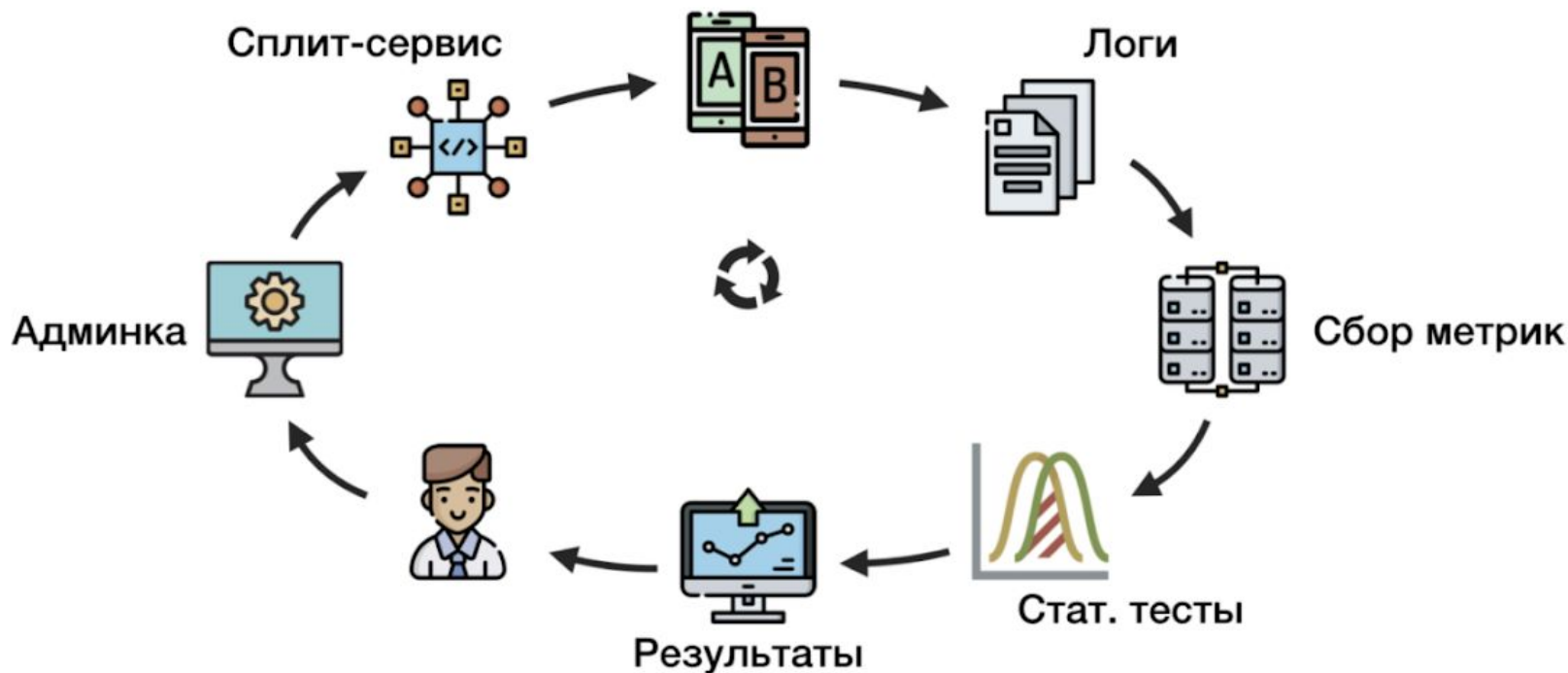


## Когда можно применять?

- Есть достаточно трафика
- Сильные гипотезы
- Инфраструктура/компетенции в команде



# Процесс тестирования



# Иерархия методов оценки причинности(Evidence ladder)



A/B тесты		<p>Юзеры поделены случайно на контроль и тест. Это дает достаточно точную оценку причинного эффекта от воздействия.</p> <p>Классическое A/B , Байесовское A/B</p>	
Квази-эксперимент		<p>Контроль и тест не идентичны и поделены на основе “натурального” критерия. В зависимости от качества критерия есть возможность также получить достаточно точную оценку причинного эффекта от воздействия.</p> <p>Differences- in-differences и другие</p>	
Искусственный контроль (counterfactuals)		<p>Контрольной группы не существует , вместо нее создается “синтетический бейзлайн” с помощью предиктивной модели, описывающей “ситуацию”, которая была бы без воздействия (counterfactual)</p> <p>Синтетический контроль , Causal Impact</p>	
Описательный анализ		<p>Нет контрольной группы, нет запущенного эксперимента. Тренды и инсайты рассматриваются с предостережением, что корреляция не равняется причинно - следственной связи</p>	

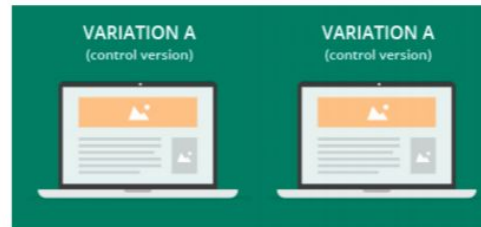
## A/A-эксперименты

Принцип схож с A/B тестами, только в данном случае обе группы видят одинаковые версии продукта

A/B



A/A



Для чего нужен данный метод ?

- Проверить качество системы сплитования пользователей.
- Выбрать гомогенные группы.
- Часто используется для симуляций.

## MVT(мультвариативные тесты)

Между собой сравниваются более двух вариантов (первая/вторая/п группа...) чтобы определить какой из вариантов эффективнее.

То же самое что и A/B только сравнивается от 2 - можно протестировать много вариантов за схожий промежуток времени.

За преимущества проверки большего числа вариантов придется заплатить объемом трафика и в некоторых случаях точностью тестов.

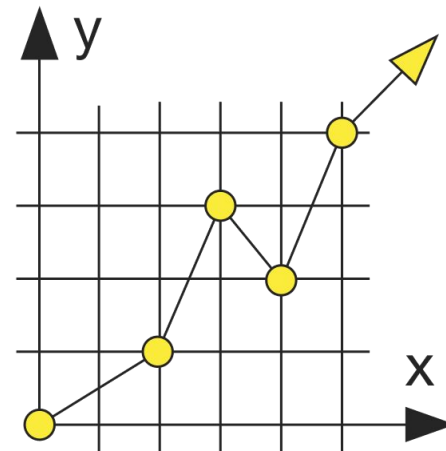


## Ухудшающие А/В-тесты.

Принцип такой же как и в классическом А/В , только в данном случае в тестовой группе делается намеренное продуктивное ухудшение .

Суть данного подхода состоит в том, чтобы завалидировать существование зависимостей метрик от определенных параметров продукта для дальнейшей генерации гипотез.

1. Нужен для предотвращения или подтверждения дорогой разработки
2. Позволяет понять чувствительность ключевых метрик
3. Можно быстро посмотреть а есть ли зависимость в принципе: например, влияет ли замедление сайта/приложения на ключевые метрики или нет.





# Чек-лист при проведении A/B теста

## Формирование и приоритезация гипотез

- ☐ Составление и приоритезация бэклога экспериментов

## Проектирования дизайна эксперимента для проверки гипотез

- ☐ Сформулирована цель и гипотеза
- ☐ Выбраны метрики и длительность эксперимента
- ☐ Проверена инфраструктура
- ☐ Все ответственные за эксперимент оповещены не позднее чем за один день до начала эксперимента

## Запуск теста

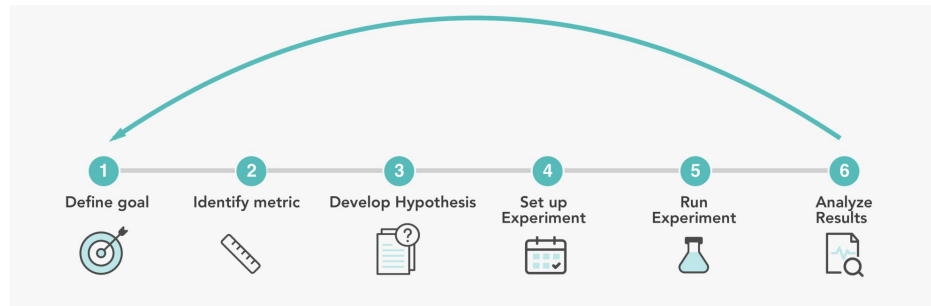
- ☐ Выполнено корректное сплитование юзеров
- ☐ Все стейкхолдеры знают о запуске теста и его деталях
- ☐ Создан дашборд

## Сбор и анализ данных. Интерпретация результата

- ☐ Не делаются финальные выводы на основе промежуточных результатов
- ☐ Посчитаны результаты и сделано описание
- ☐ Коллеги оповещены, когда рез-т готов

## Принятия решения на основе результатов теста

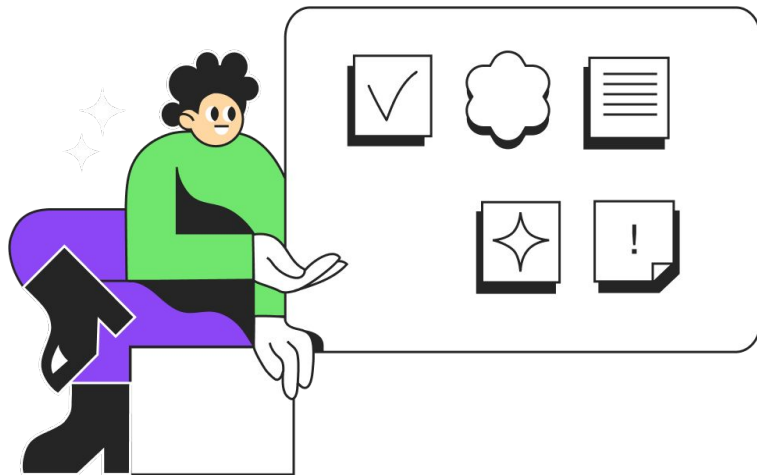
- ☐ Принято решение по результатам теста





## На этом уроке мы разобрали

- 💡 Рассмотрели какие виды экспериментов существуют
- 💡 Обсудили как понять можно ли и нужно ли вам проводить А/Б-тест
- 💡 Посмотрели как выглядит процесс А/В тестирования





## На следующем уроке мы рассмотрим:

- поговорим о настройках а/б тестов в веб приложениях
- запустим и с нуля настроим а/б тест в Google Optimize