

**A/B тесты**

- Что же такое A/B тесты
- Важные моменты
- достоинства и недостатки A/B тестов

Допустим, мы работаем в интернет-издании. Каждый день к нам заходят пользователи и смотрят наш контент. Наша задача рекомендовать пользователю максимально релевантные статьи, чтобы пользователь больше времени проводил на нашем сайте, возвращался в дальнейшем.

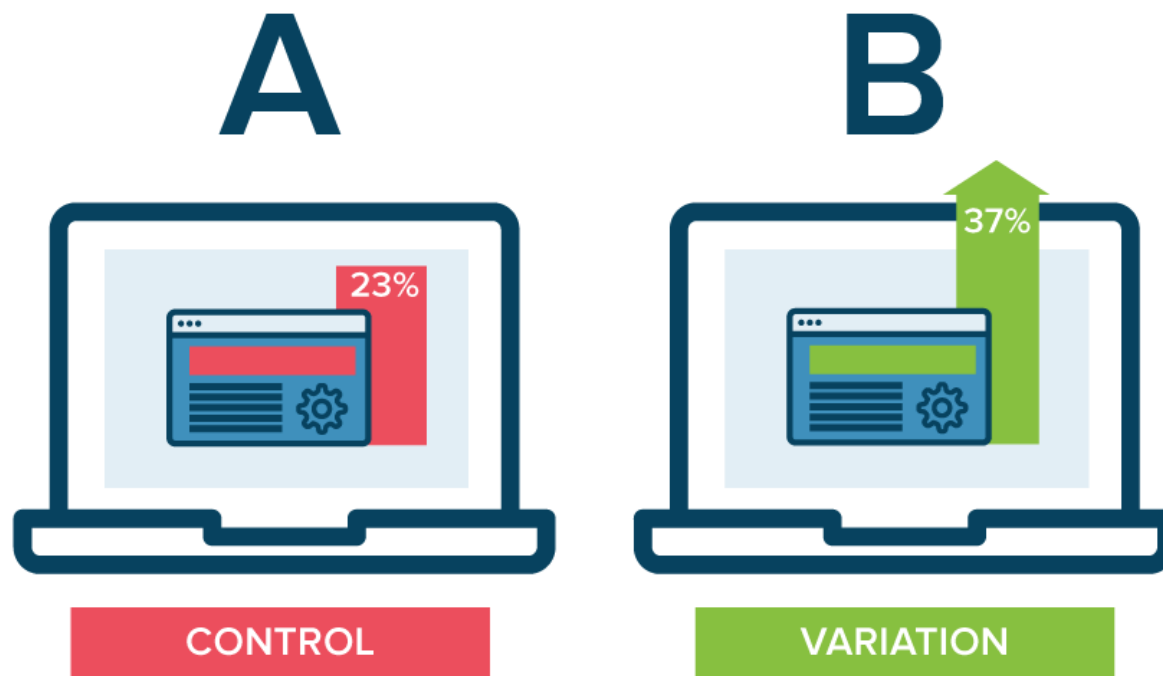
- Ранее у нас уже была какая-то модель рекомендации, но у нас появилась новая идея, как сделать рекомендации лучше.
- Если мы ее внедрим, как количественно оценить, что рекомендации стали лучше?
- Именно этим и занимается **A/B - тестирование**.

# Обобщенный алгоритм A/B теста

Коротко идею A/B - тестирования можно описать следующим образом:

- Наш трафик делится на две группы(может и больше), называемый *контрольной* и *экспериментальной*
- *Контрольная* группа получает старые рекомендации
- *Экспериментальная* группа получает новые рекомендации
- После проведения A/B - теста зафиксировать измеряемую нами метрику в контрольной и экспериментальной группах и сделать вывод, есть ли **статистически значимые(!)** отличия.

# Обобщенный алгоритм A/B теста



# Для чего нужны А/В - тесты

- Создание любого продукта — непрерывный процесс поиска улучшений, которые нужно фиксировать.
- Каждая гипотеза или идея нововведения требует проверки — иначе наш продукт может уйти не туда.
- Ожидаемое улучшение не всегда на самом деле улучшает продукт — новый функционал может быть никому не нужен или реализация портит всю затею.

- Основные требования к корректному проведению **A/B - теста**:
  - Наблюдения (пользователи) независимы
  - группы пользователей эквивалентны (из одной ГС)

Последнее может нарушаться по разным причинам, например:

- данные приходят в систему аналитики корректно.
- инструмент «деления» нашего трафика работает ошибочно

Чтобы удостовериться, что группы пользователей действительно эквивалентны (из одной ГС), и нет никаких "человеческих ошибок" при подготовке данных необходимо провести так называемый **А/А - тест**:

- Принцип работы точно такой же, как и при А/В - тесте, только и контрольная и экспериментальная группы видят одно и то же.
- Если ключевые для нас метрики не имеют статистически значимых различий, считаем, что группы пользователей эквивалентны
- После проведения А/А - теста. Составы групп остаются теми же для проведения А/В -теста



# Опережающие метрики

- Важнейшим этапом является выбор метрики, изменение которой вы хотите получить. Не все метрики поворотны и чувствительны к небольшим изменениям. Воздействовать на метрики аналитику, дата-саентисту не всегда просто. Результаты зависят от множества факторов и работы огромного количества людей.
- Важно понимать, на какие метрики ваше изменение способно повлиять
- **Основные метрики** - ключевые бизнес-метрики компании, которые отражают здоровье бизнеса (*Выручка, валовая прибыль, средний чек, конверсия в покупку*)
- **опережающей(или прокси) метрикой** называют метрику, изменение которого напрямую влияет на основную метрику компании или подразделения.  
(*Продолжительность сессии, количество посещений, конверсия*)

Мы оцениваем эффективность изменения дизайна меню товаров интернет-магазина фермерских продуктов. Изменения(если они действительно есть) на основных метриках сразу увидеть у нас не получится. Какие опережающие метрики можете предложить?