

Ускорение и оптимизация А/В-тестов

Артур Сапрыкин

Цель модуля

Обсудить все возможные способы
и техники по ускорению тестирования.
Начинать будем с простых,
но эффективных!

Задачи модуля

- ✓ Рассмотрим, как можно снизить число экспериментов, необходимых для проверки гипотез при прочих равных
- ✓ Вы узнаете, как стандартизировать принятие решений, чтобы не терять время при появлении новых результатов
- ✓ Заглянем под капот некоторых продвинутых методов снижения дисперсии в данных

Ускорение и оптимизация A/B-тестов

Простые методы ускорения A/B-тестов

Цель урока

Начнём с методов, позволяющих ускорить время принятия решения при проверке гипотез. В этом уроке мы обсудим простые, но эффективные способы ускорить тестирования, с которых стоит начинать в первую очередь.

Задачи урока

- ✓ Поговорим про ухудшающие тесты. Посмотрим на примеры
- ✓ Разберём аргументы за и против ухудшающих тестов
- ✓ Затронем тему количества времени, расходуемого на каждый этап цикла тестирования

Сколько времени занимает А/В-тест?

В теории:

Сам
А/В-тест

9–12 дней

Анализ
результата

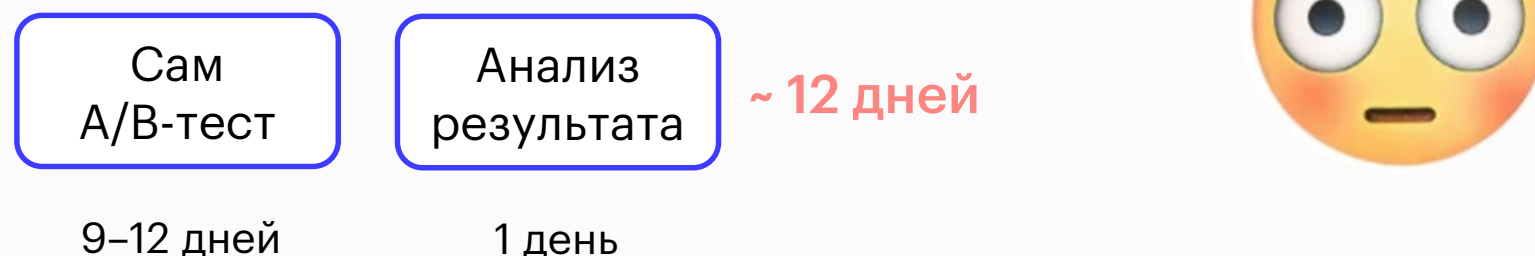
1 день

~ 12 дней

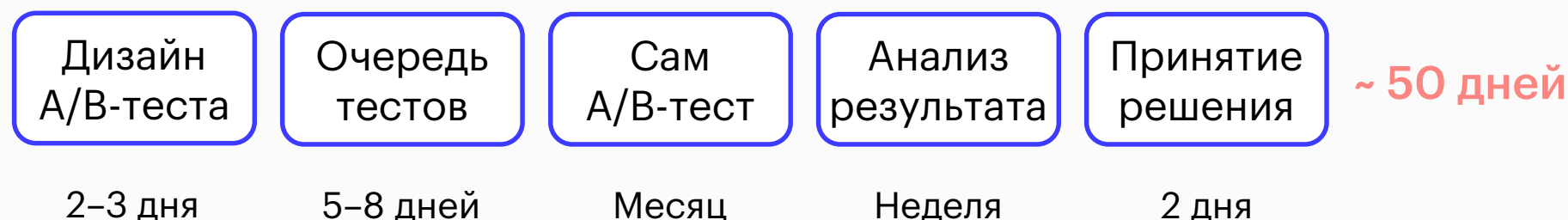


Сколько времени занимает А/В-тест?

В теории:

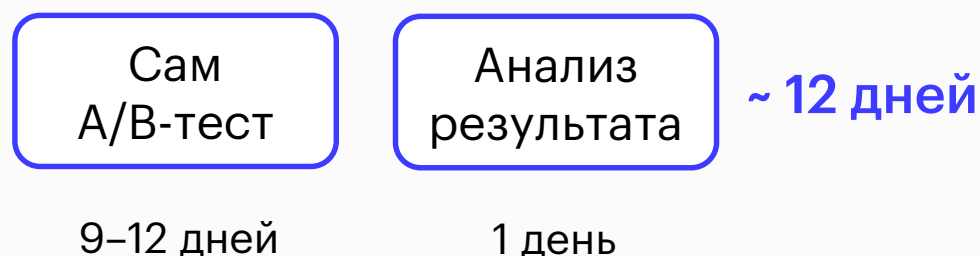


В реальной жизни:



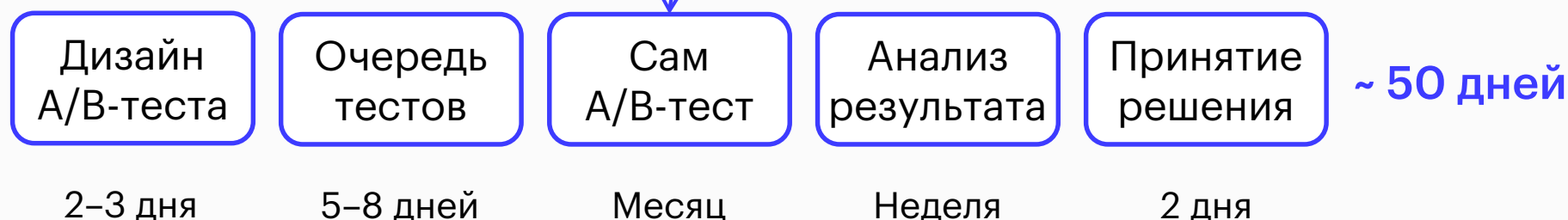
Сколько времени занимает А/В-тест?

В теории:



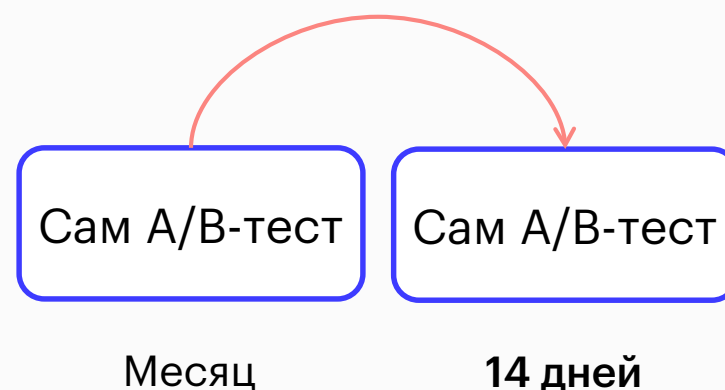
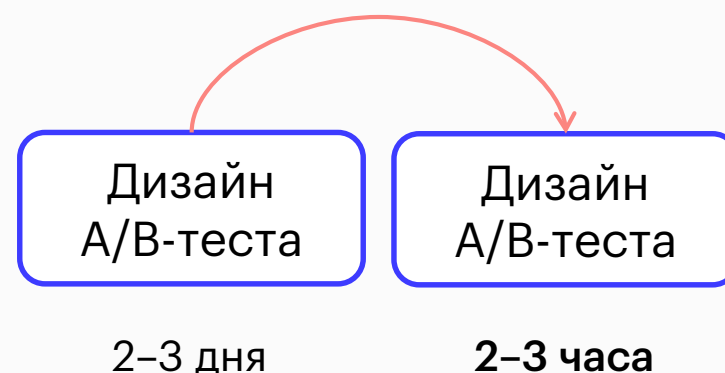
Получается, чтобы ускорить А/В-тестирование, недостаточно просто как можно быстрее собрать необходимое количество данных, надо сделать ещё...

В реальной жизни:



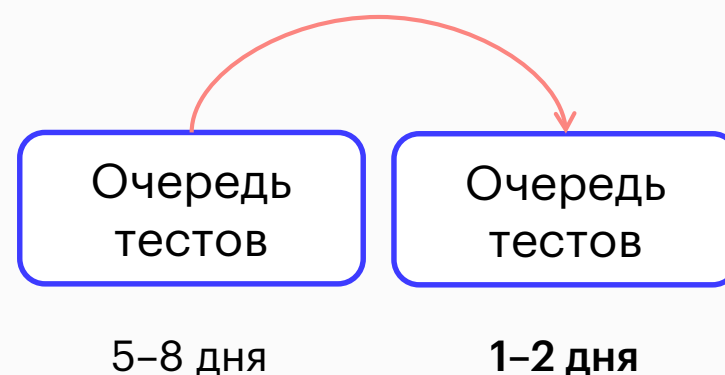
Единый стандарт дизайна теста

- ✓ Правила выбора:
 - $\alpha = 5\%$, power = 80 %, количество групп ≤ 3 , метрик ≤ 3
 - Такие α и power вполне себя хорошо показали. Надо начинать с них
- ✓ Стандартные метрики направлений (поиск, корзина)
- ✓ Длительность теста — 2–3 недели
- ✓ Валидация A/B-теста до запуска!



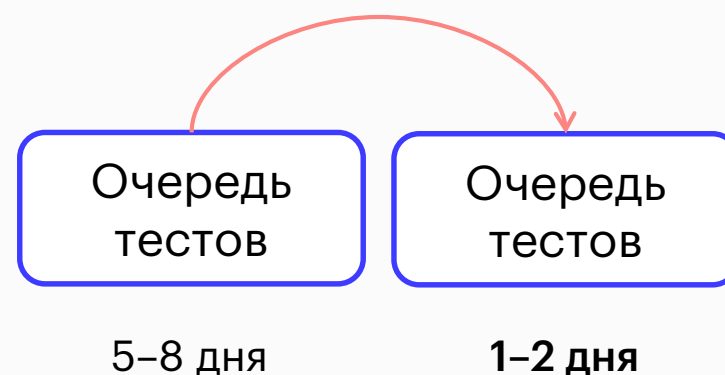
Расписание/календарь A/B-тестов

- ✓ Мониторится очередь экспериментов
- ✓ Можно забронировать нужный бакет пользователей
- ✓ Удобно приоритезировать и планировать запуски
- ✓ Необходимо, если в компании много отделов



Расписание/календарь A/B-тестов

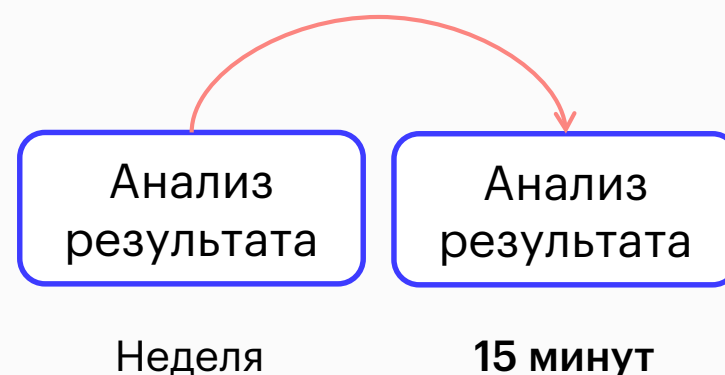
- ✓ Мониторится очередь экспериментов
- ✓ Можно забронировать нужный бакет пользователей
- ✓ Удобно приоритезировать и планировать запуски
- ✓ Необходимо, если в компании много отделов



Теперь время сокращено, а запуски тестов вы легко можете спланировать, посмотрев в календарь.

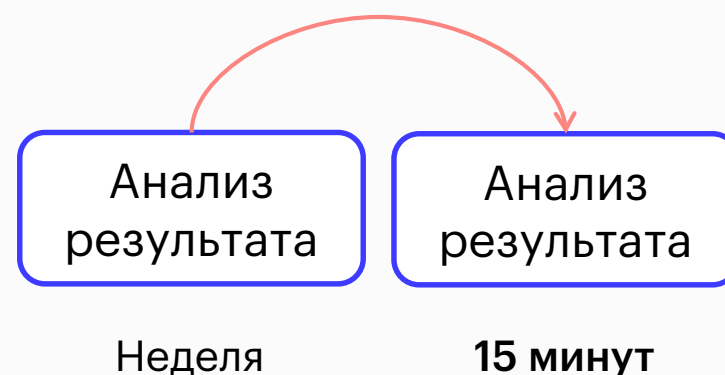
Автоматизированный анализ/расчёт результатов тестов

- ✓ Автоматическое преобразование / чистка данных
- ✓ Автоматическое применение статистических критериев
- ✓ Единообразие подсчёта метрик во всей компании
- ✓ Пайплайны для расчёта всех метрик, возможность видеть этапы (AirFlow)
- ✓ Необходимо, если в компании много отделов



Автоматизированный анализ/расчёт результатов тестов

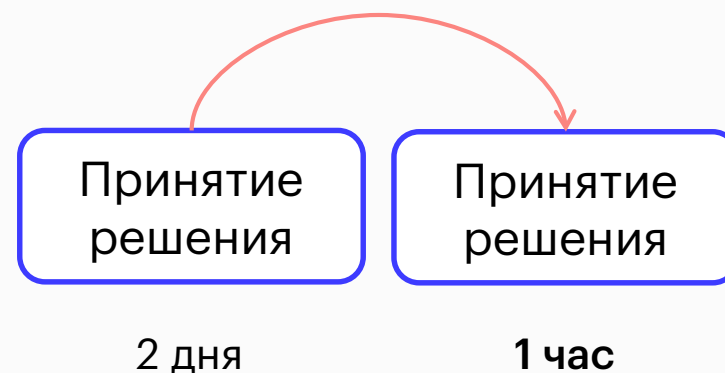
- ✓ Автоматическое преобразование / чистка данных
- ✓ Автоматическое применение статистических критериев
- ✓ Единообразие подсчёта метрик во всей компании
- ✓ Пайплайны для расчёта всех метрик, возможность видеть этапы (AirFlow)
- ✓ Необходимо, если в компании много отделов



Теперь этот этап происходит быстро, а время аналитиков не растрачивается зря.

Выработанная культура принятия финальных решений

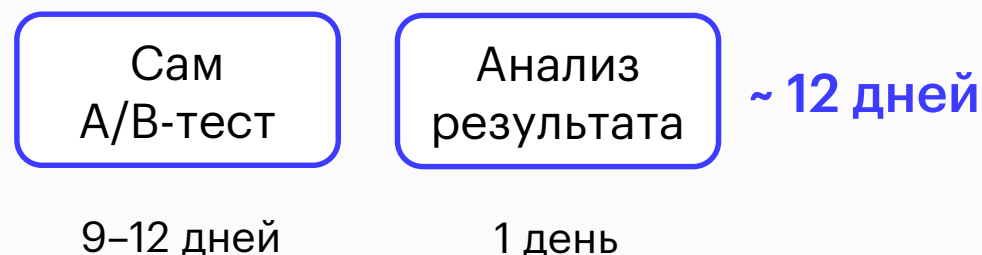
- ✓ Заранее планировать сроки.
Не ждать ещё, пока всё-таки получится значимость
- ✓ Не затягивайте с выкаткой
или отказом от фичи
- ✓ Если вдруг гипотеза не дала
результат — не искать другие
метрики, где результат
всё-таки успешный



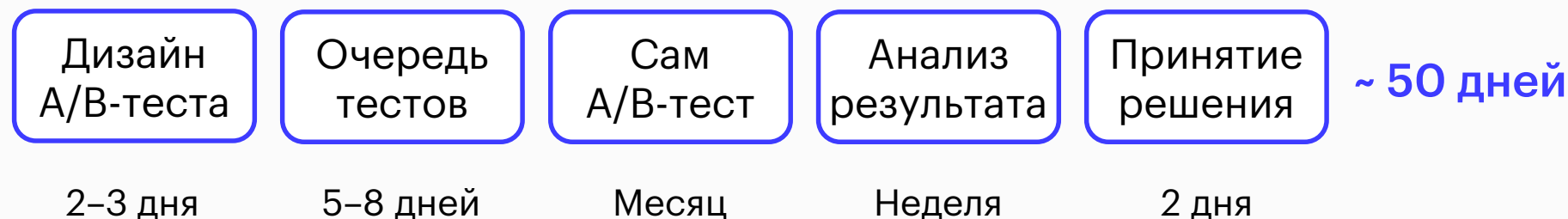
Теперь и этот этап ускорен, а вы не мусолите неоднозначные ситуации, а просто переходите к новым гипотезам.

Сколько времени занимает А/В-тест?

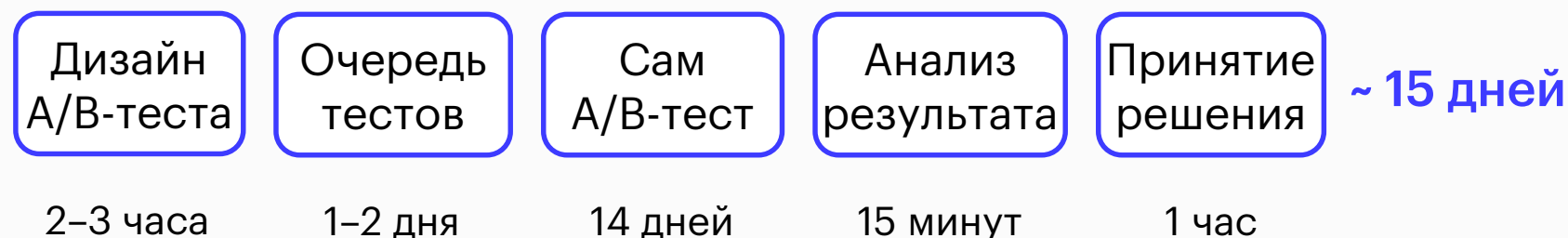
В теории:



В реальной жизни:



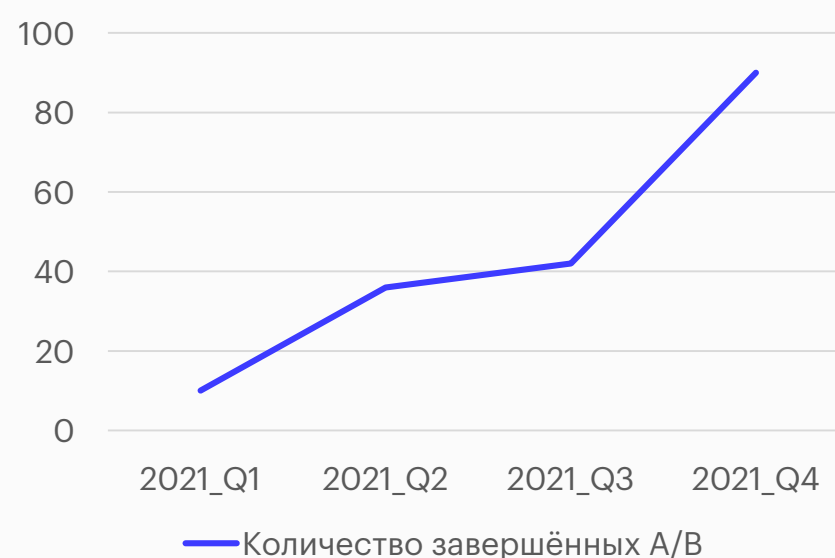
В реальной жизни:



Кейс от Delivery Club: ускорение тестирования

Наибольший вклад внесли:

- слои
- авторасчёт
- линейаризация + прокси

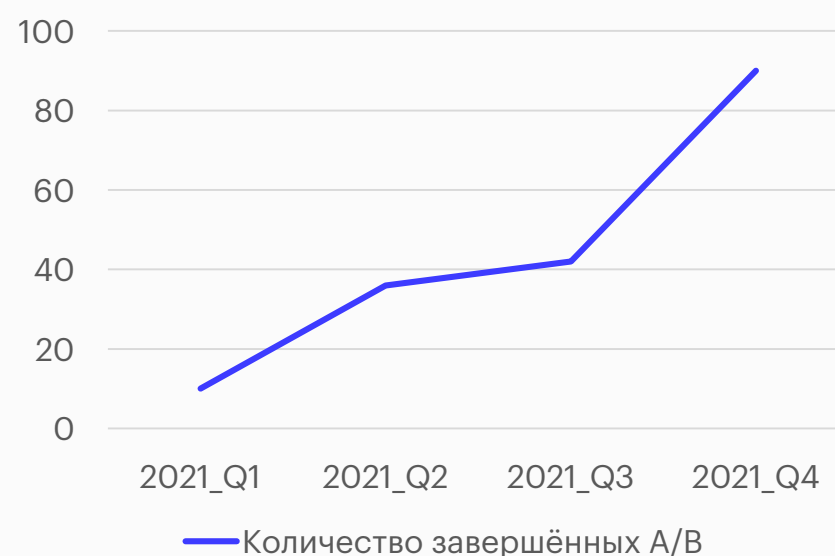


Кейс от Delivery Club: ускорение тестирования

Наибольший вклад внесли:

- **слои**
- авторасчёт
- линейаризация + прокси

Возможно проводить несколько видов тестов на одном пользователе (бакете).

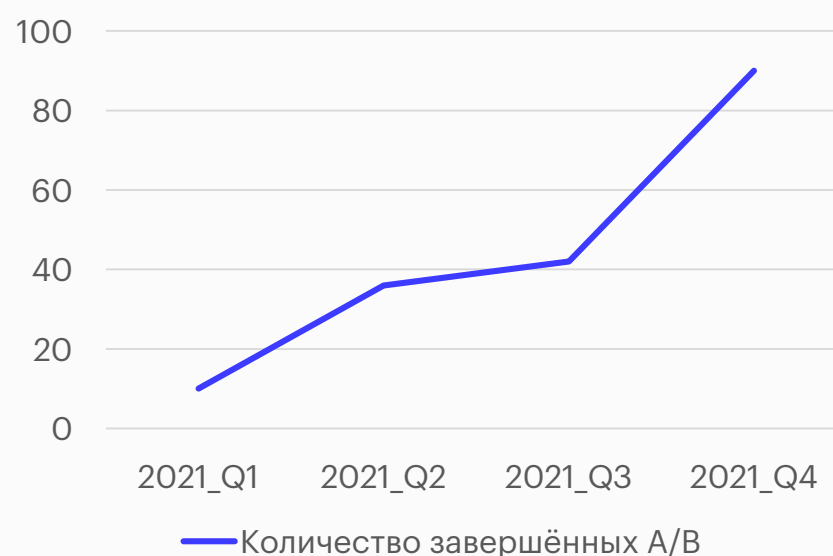


Кейс от Delivery Club: ускорение тестирования

Наибольший вклад внесли:

- слои
- авторасчёт
- линейаризация + прокси

Если столько метрик считать руками, то аналитики будут этим заниматься 24/7 — автоматизация спасает!



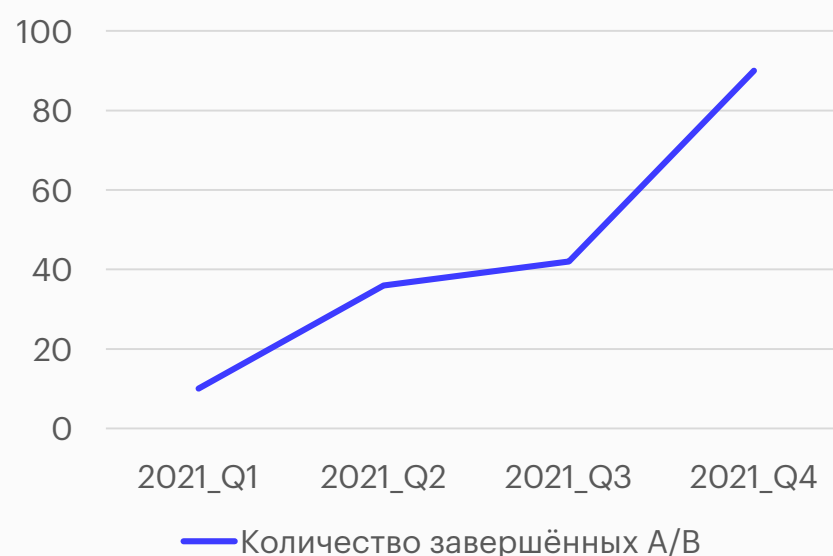
Кейс от Delivery Club: ускорение тестирования

Наибольший вклад внесли:

- слои
- авторасчёт
- **линеаризация + прокси**

Способ увеличения
чувствительности метрик.

Мы про это ещё будем
говорить дальше.



Итоги и выводы урока

Сильно ускорить ваши тесты можно на разных этапах. Если в вашей компании пока нет A/B-тестов, то стоит начать со стандартизации тестов по следующим 4 шагам, при этом полезно понимать, в какой примерно момент времени стоит этот этап реализовывать.

Итоги и выводы урока

- ✓ **Создайте единый подход к тестированию** (10–20 тестов/квартал): α , power, количество групп, дизайн
- ✓ **Автоматизируйте расчёты!** (20–40 тестов/квартал): стандартные метрики, авторасчёт, календарь
- ✓ **Настройте массовое A/B-тестирование** (40–200 тестов/квартал): слои, культура принятия решений
- ✓ **Ускорение метрик!** (200+ тестов/квартал): линализация, CUPED, трансформация метрик

Про последний шаг мы ещё поговорим подробнее в этом модуле!