



Как делать выводы из данных

Tinkoff.ru

Апрель, 2021

Не забыть представиться 😊





Вступление: зачем это все

Как анализировать данные

Как делать выводы

Как считать деньги

= как связывать данные и жизнь

Зачем тебе аналитика?



Аналитика – это не просто «посчитай по факту».



1. Принять решение



2. Найти зоны роста/успеха в процессе, исследовать на оптимальность



3. Понять, что реально происходит в продукте

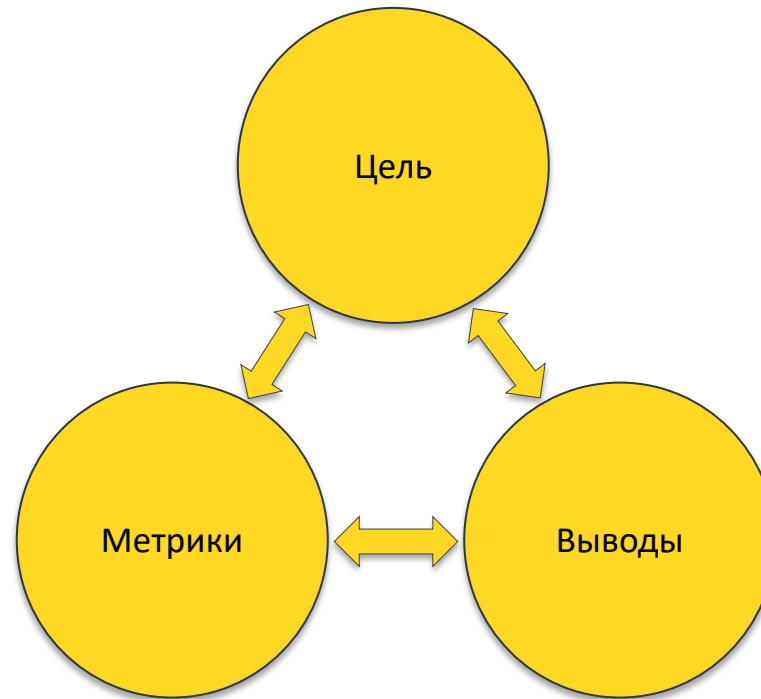


4. Мониторинг, «рука на пульсе»



5. Прогнозирование, планирование

Зачем тебе аналитика?



Переходим к метрикам



Пример 1. Много плохих отзывов

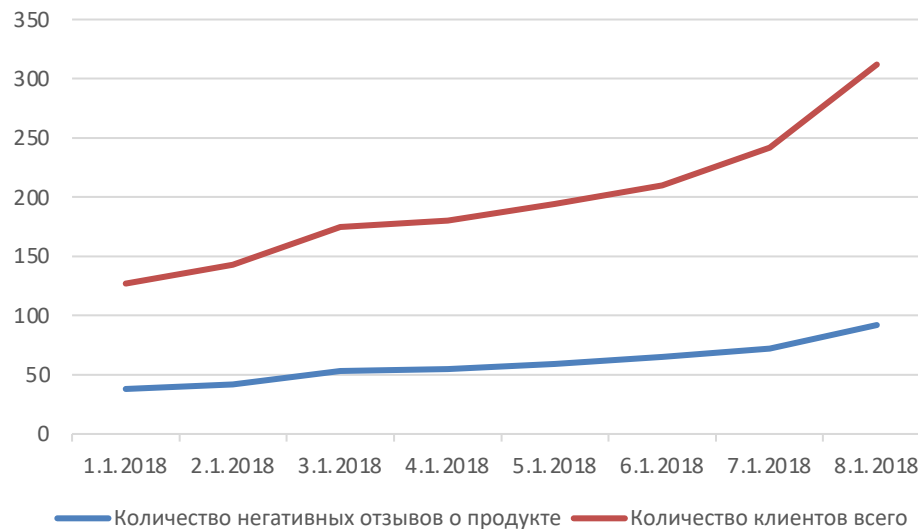


Продукт испортился?

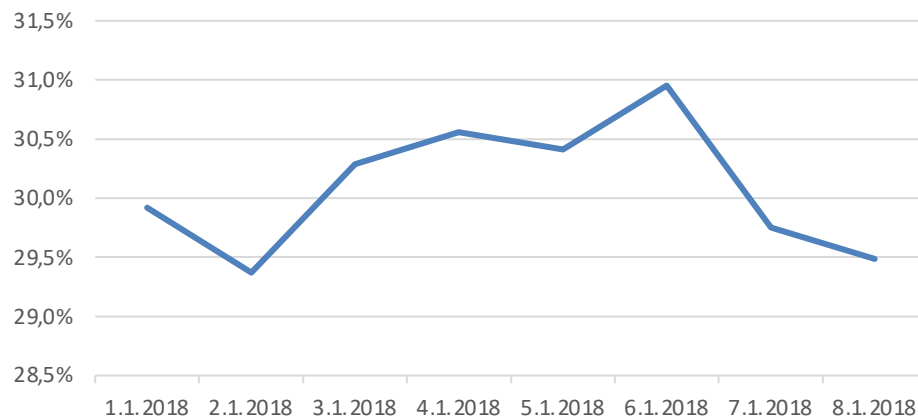
Доля вместо абсолютных значений



Количество негативов и клиентов всего



Процент негативных отзывов



Пример 2. Нужно ли исправлять работу процесса?

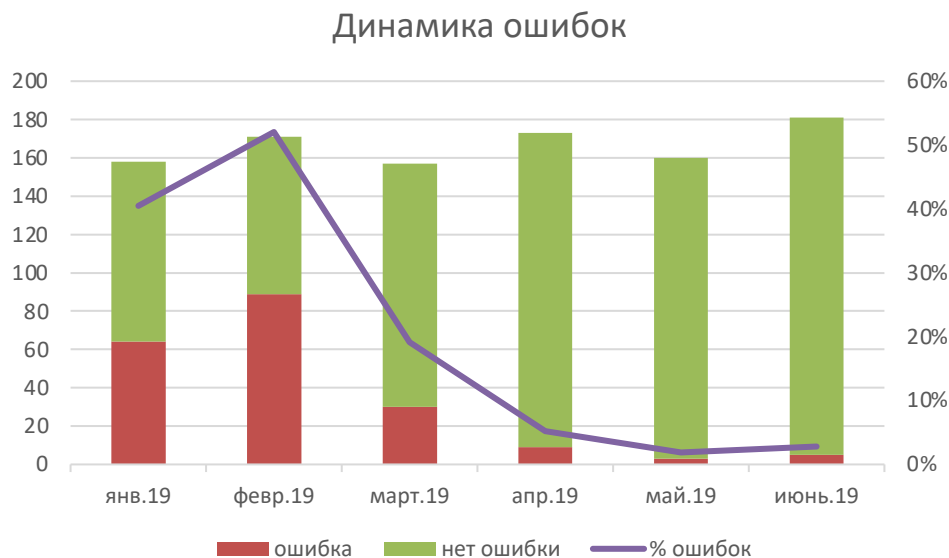


Посчитали кол-во ошибок в запусках процесса за прошлые полгода.

- Всего запусков: 1000
- Из них с ошибкой: 200
- 20% - много или мало?

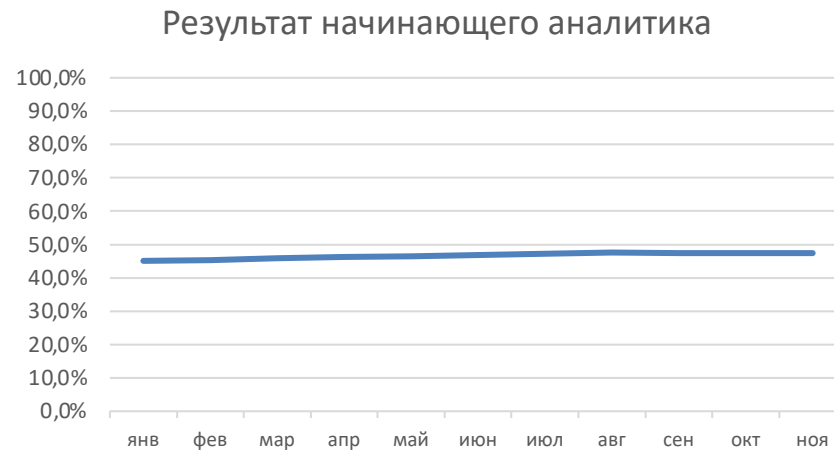
Посчитали кол-во ошибок в запусках процесса **за прошлые полгода.**

- Всего запусков: 1000
- Из них с ошибкой: 200
- 20% - много или мало?



просто результат -> стабильность во времени

Пример 3 (жизненный)



Помимо самого графика и тренда на нем – смотрите на:

- Масштаб
- Числа
- Статзначимость (объем выборки)

Вспомнили предыдущую лекцию



База рассмотрения метрик:

- Статика / динамика
- Абсолютные / относительные
- Правильная визуализация



Если конверсия была 20%, а стала 10%
- это значит что конверсия уменьшилась
на 50 процентов
и
на 10 процентных пунктов 😊

Метрики продукта и метрики роста



Сравниваем две булочные

	Булочная А	Булочная Б
Прибыль за последний год	2 млн	4 млн
Кол-во клиентов	1000	5000
Средний чек за последний год	200р	300р

Вы – инвестор. Какую из них купите?



Метрики продукта и метрики роста

Сравниваем две булочные

	Булочная А	Булочная Б
Прибыль за последний год	2 млн	4 млн
Кол-во клиентов	1000	5000
Средний чек за последний год	200р	300р
% прибыли от стоимости булочной	30%	5%
% возвращающихся клиентов	20%	6%
Средний чек за последний месяц	300р	200р





Метрики продукта и метрики роста

Сравниваем две булочные

	Булочная А	Булочная Б
Прибыль за последний год	2 млн	4 млн
Кол-во клиентов	1000	5000
Средний чек за последний год	200р	300р
% прибыли от стоимости булочной	30%	5%
% возвращающихся клиентов	20%	6%
Средний чек за последний месяц	300р	200р



Метрики продукта и метрики роста: сравнение



Привлечение * метрики продукта = результат

Пример: два канала оформления карты

	Первый	Второй	
За прошлый месяц выдано карт	1000	2000	← Сравнить неправильно
Конверсия из лида в выдачу	3%	20%	← Сравнить правильно – это характеристика канала
Потенциал объема лидов/мес	2млн	1млн	← Хорошо бы это учитывать
Потенциал результата	60K	200K	← Сравнить правильно – это результат канала



Булочная:

Выручка с продаж	Расходы на производство
100.000	100.000



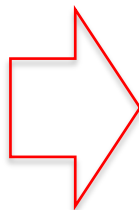
Булочная:

	Выручка с продаж	Расходы на производство
Всего:	100.000	100.000
Булочки	25.000	15.000
Торты	15.000	15.000
Кофе	40.000	10.000
Котлеты	20.000	60.000

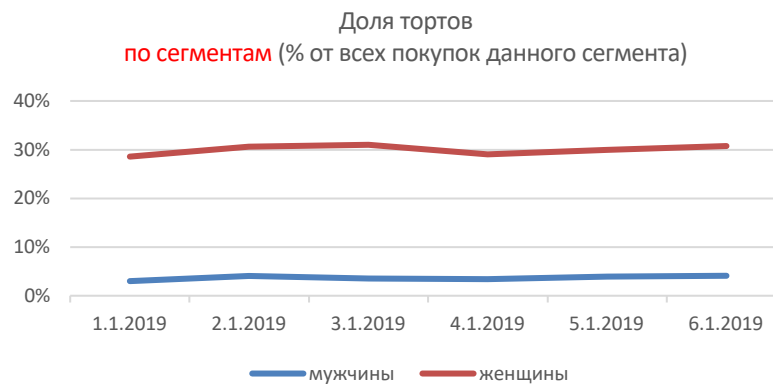
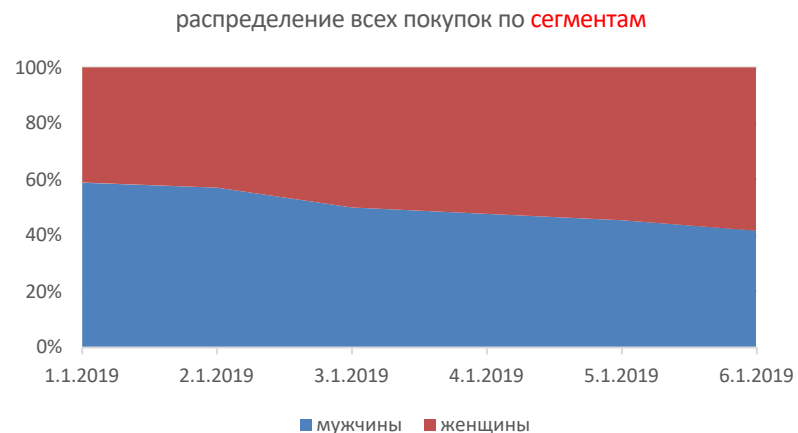
Пример на сегментацию



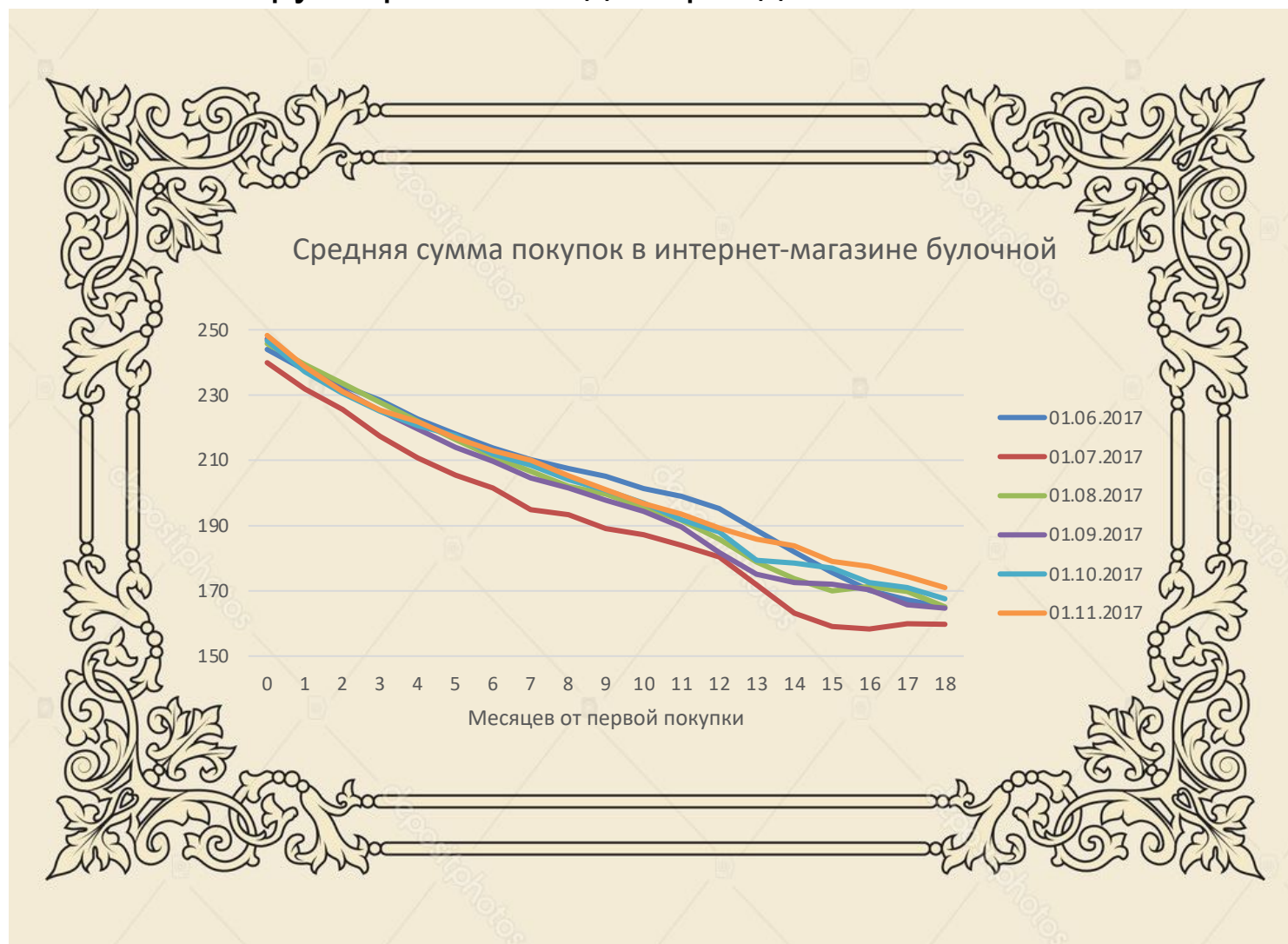
Показатель по потоку меняется...



...на самом деле это меняется распределение сегментов, а внутри сегментов все как обычно:

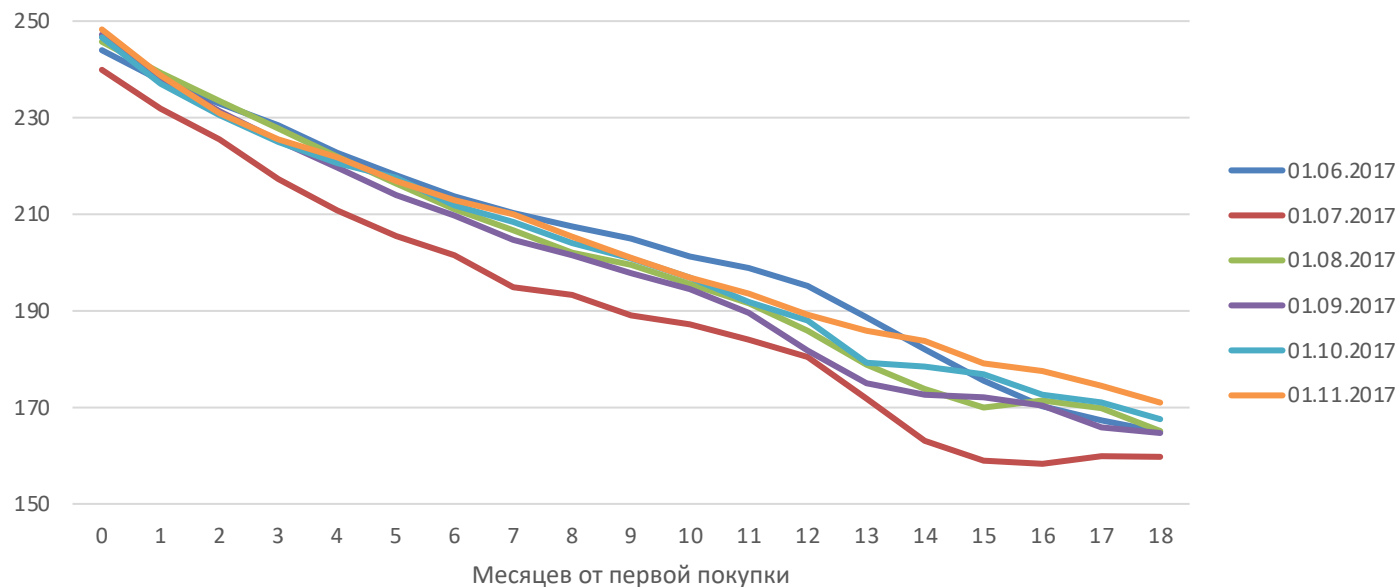


Винтажный = когортный = по поколениям
Группировка по «дате рождения»





Средняя сумма покупок в интернет-магазине булочной



Рост количества новых клиентов => средняя сумма ВСЕХ покупок растет

Новые клиенты не приходят => средняя сумма ВСЕХ покупок падает



Винтажный анализ – зачем?

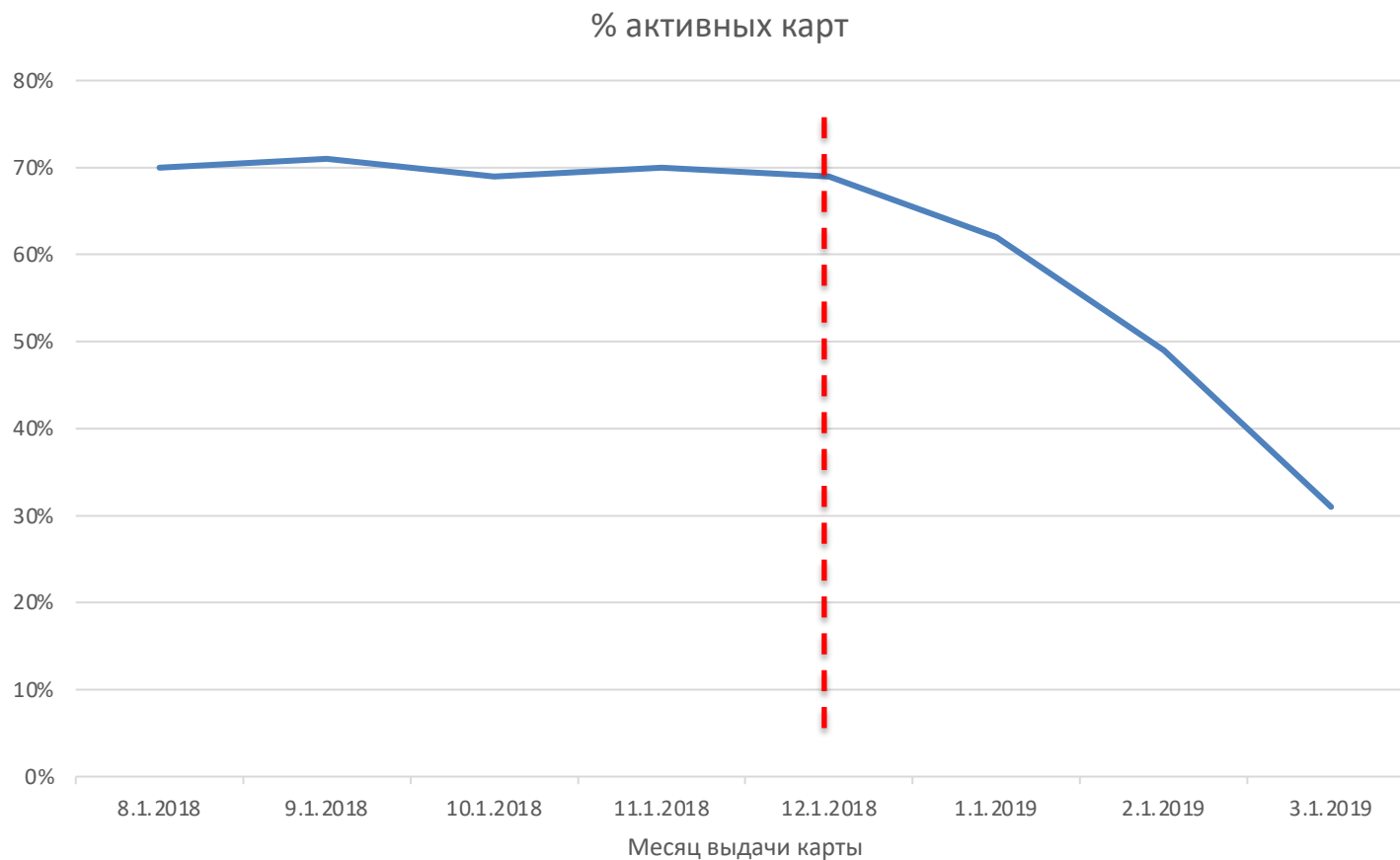
Учитывать изменение поведения со временем!

Видеть тренды. Это делается именно по винтажам, а не по ежемесячным графикам.

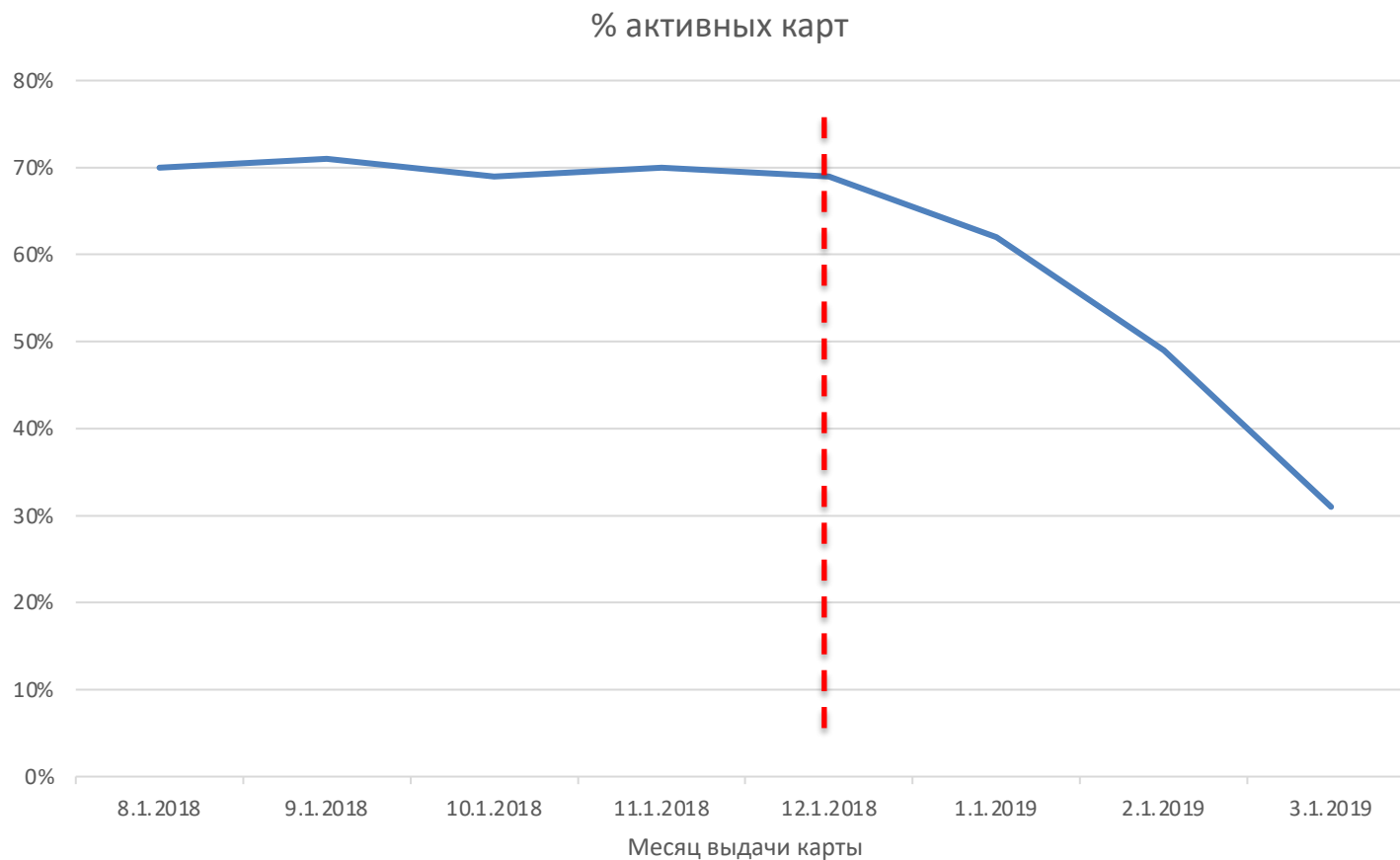


Портфель – множество клиентов разного возраста, разных характеристик, разного поведения.
Не всегда полезно мешать всех в одну кучу.

Почему конверсия падает? Сейчас апрель 2019.

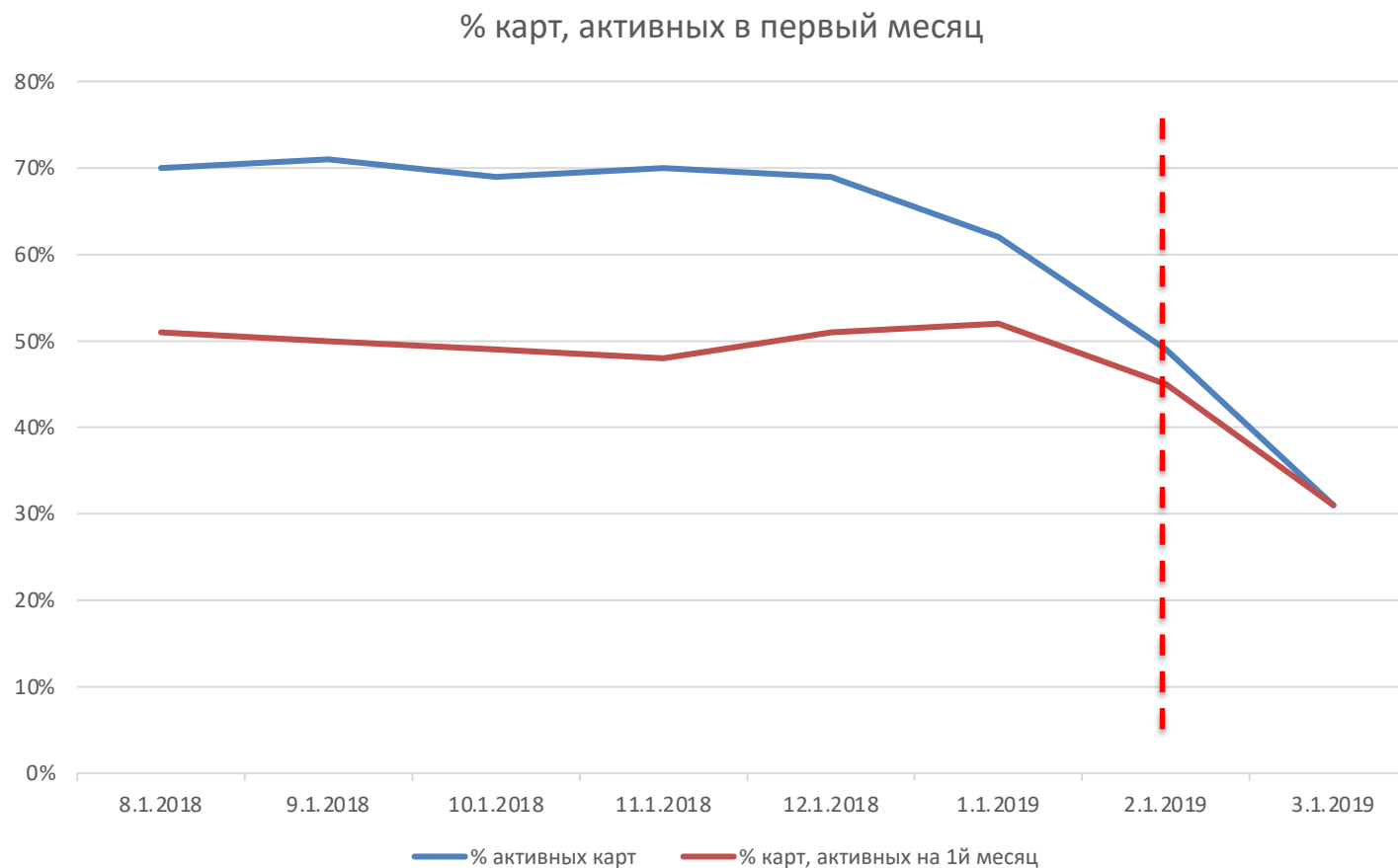


Эффект созревания



Недозрело

Эффект созревания



Отлавливаем проблемы раньше



- Полезно смотреть промежуточные метрики: конверсия за X дней
- Но: не забывать и про целевые
- Пример: смотрели только на первый этап воронки, не смотрели воронку до утилизации => потеряли деньги на привлечение

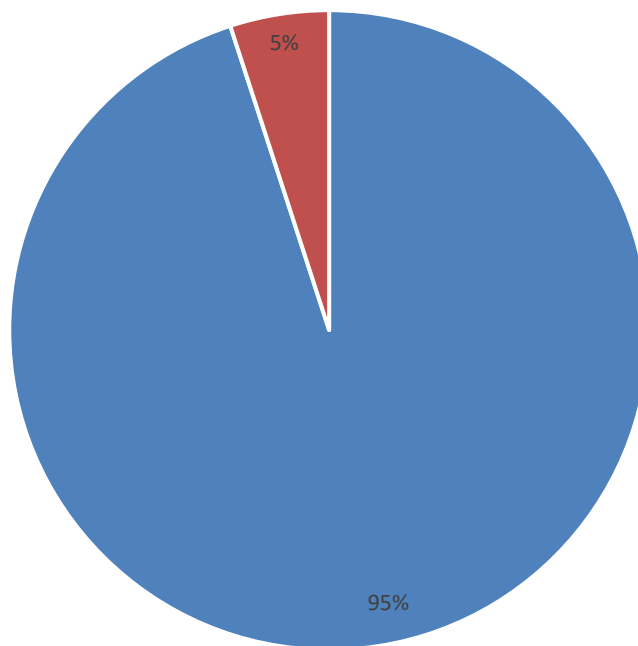


Теперь мы научились выбирать правильные метрики

Как делать выводы?



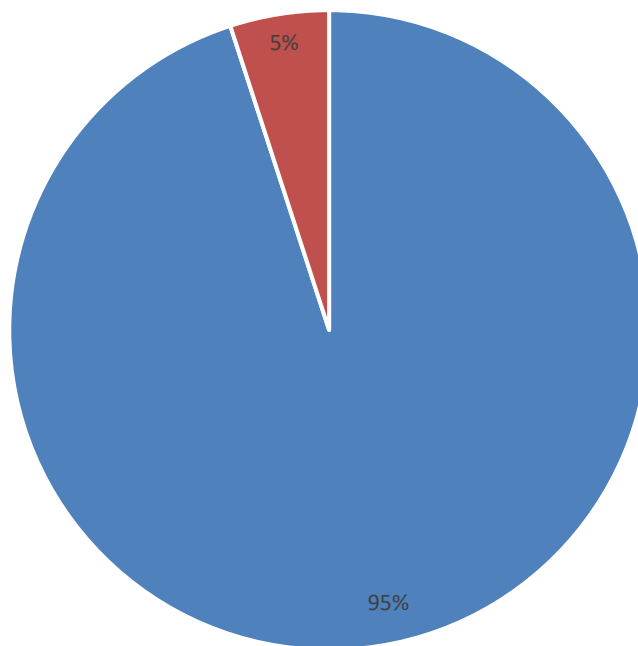
Проведен емейл-опрос среди клиентов:
какой способ бронирования квеста удобнее?



■ в интернете ■ по телефону



Проведен емайл-опрос среди клиентов:
какой способ бронирования квеста удобнее?



■ в интернете ■ по телефону

Вопрос выборки



Еще пример: ошибка выжившего



Причина или корреляция



На что обращать внимание, когда смотришь данные?

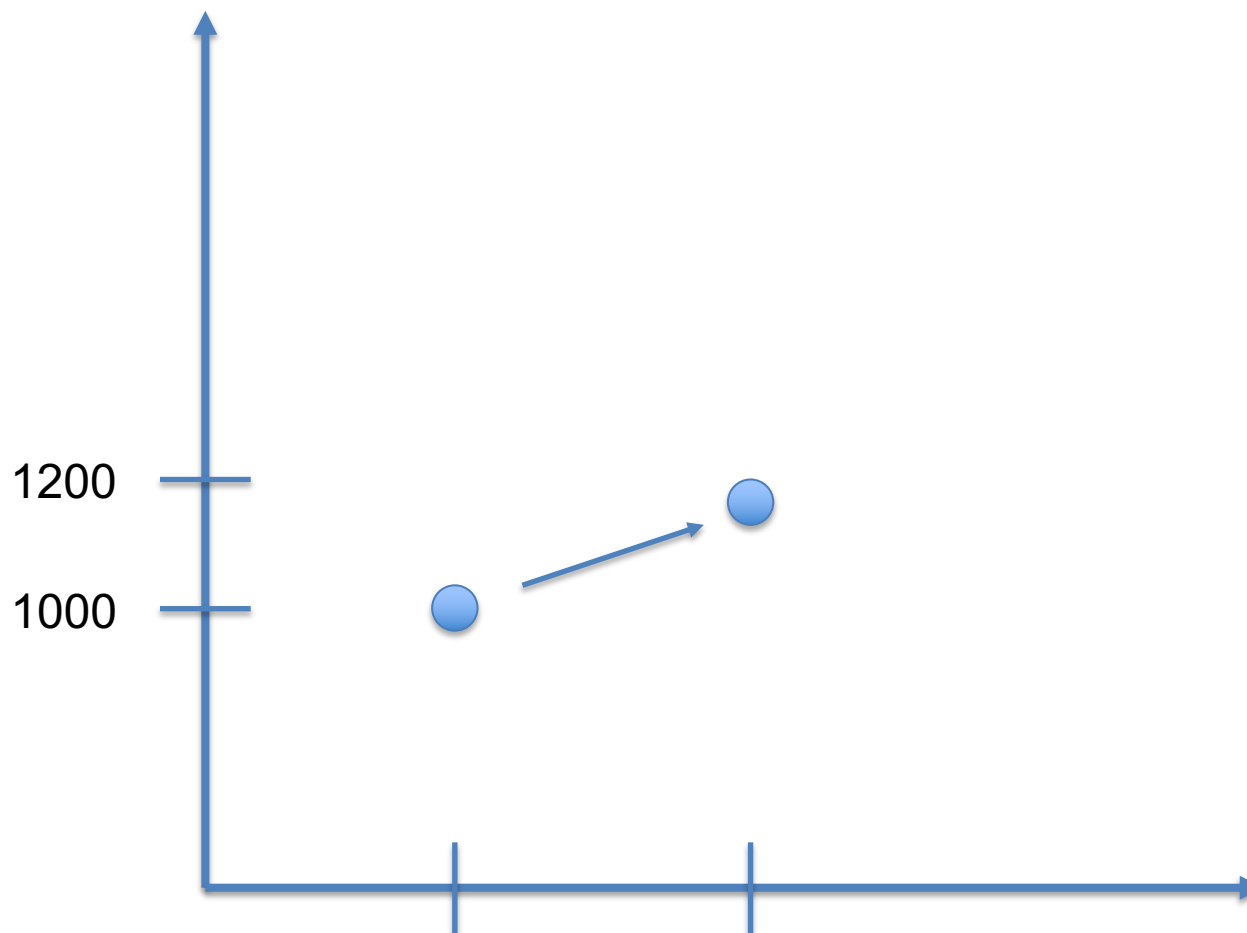


- Какая выборка?
- Почему случилось резкое изменение?
- Почему происходит плавный, но постоянный рост/падение?
- Почему цифры противоречат ожиданиям/здоровому смыслу?
- Как проверить варианты ответов на предыдущие вопросы?

Хорошо или плохо?



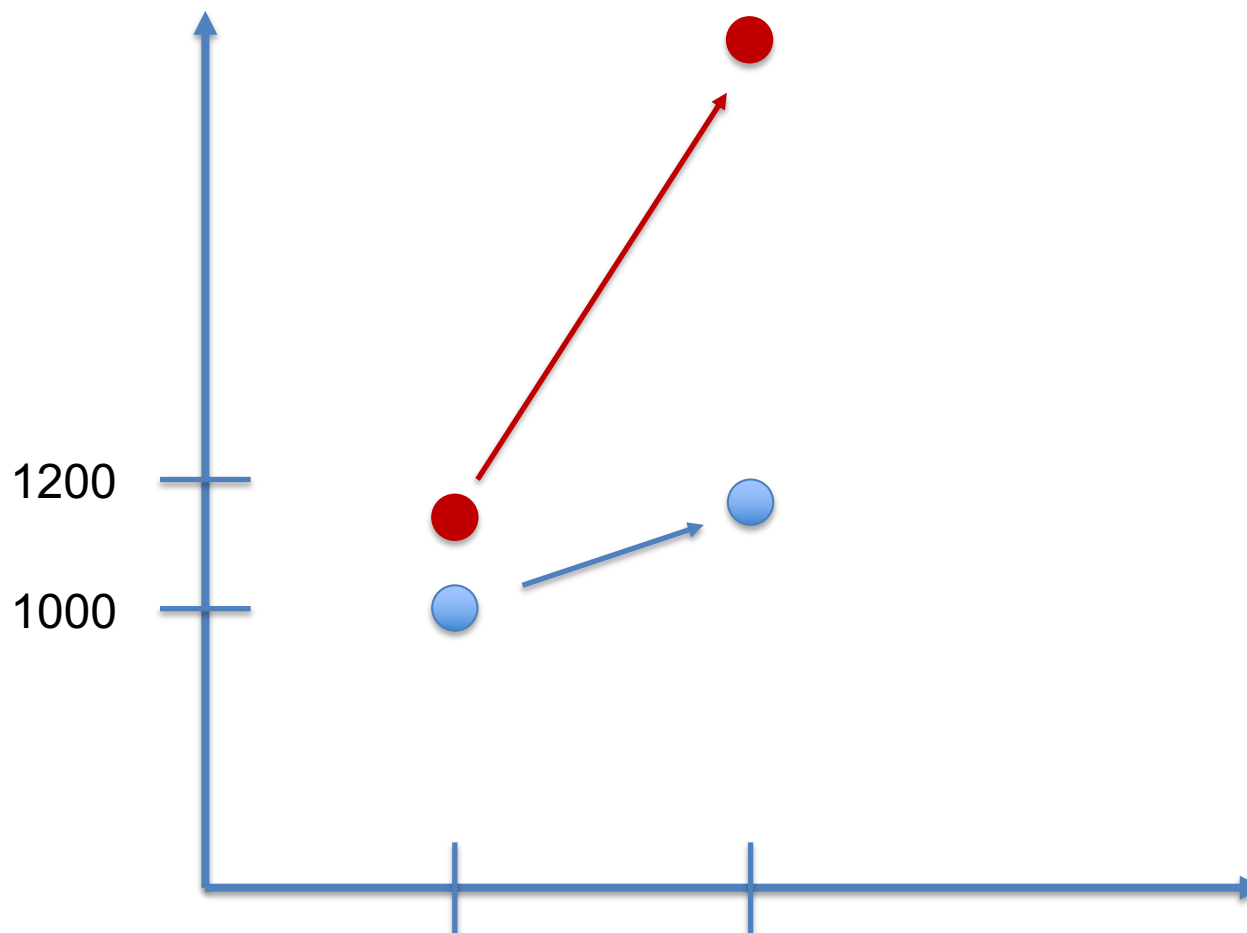
Продажи булочек в предыдущем месяце и в этом



Хорошо или плохо?



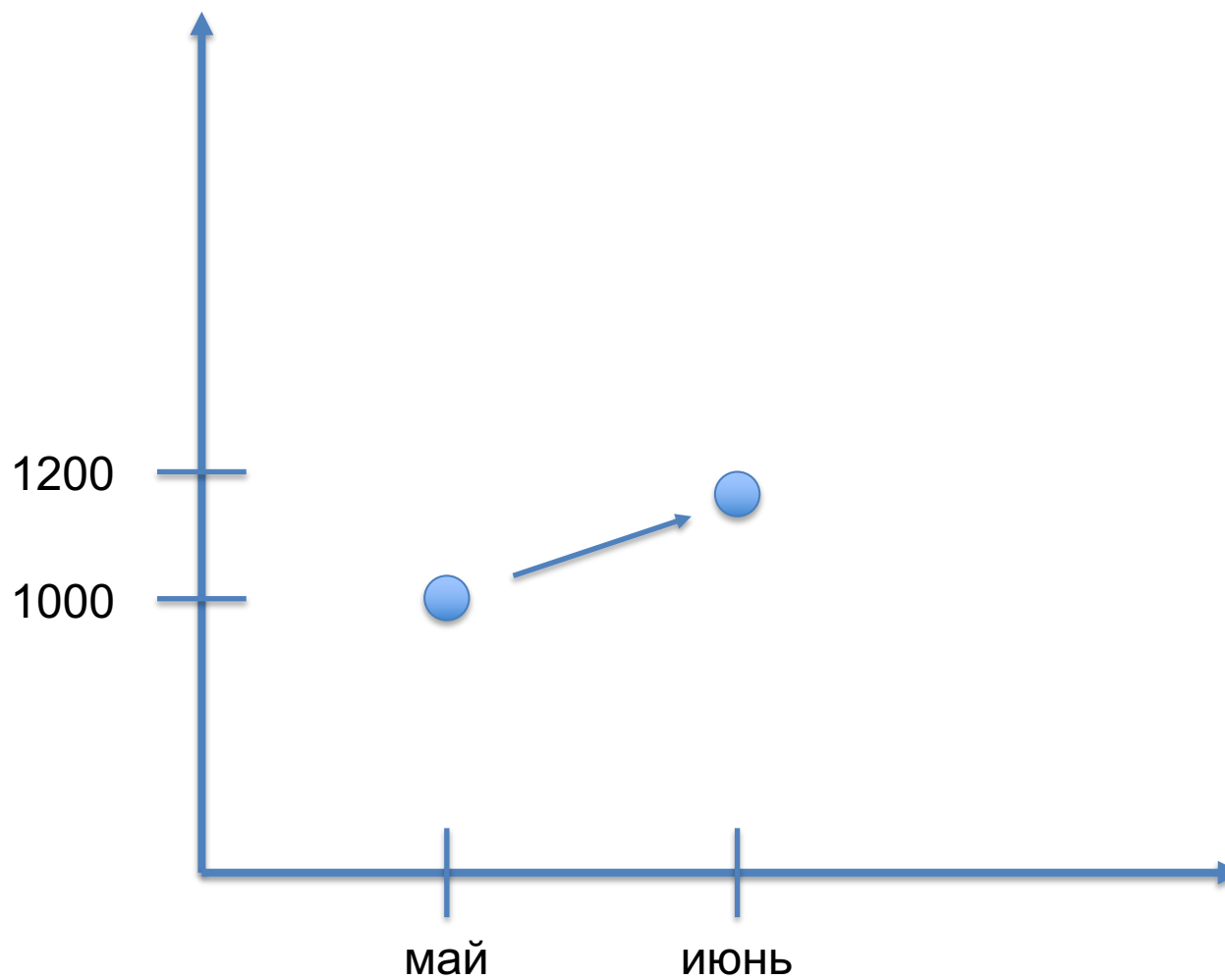
А если добавить на график кол-во посетителей сайта?



Хорошо или плохо?



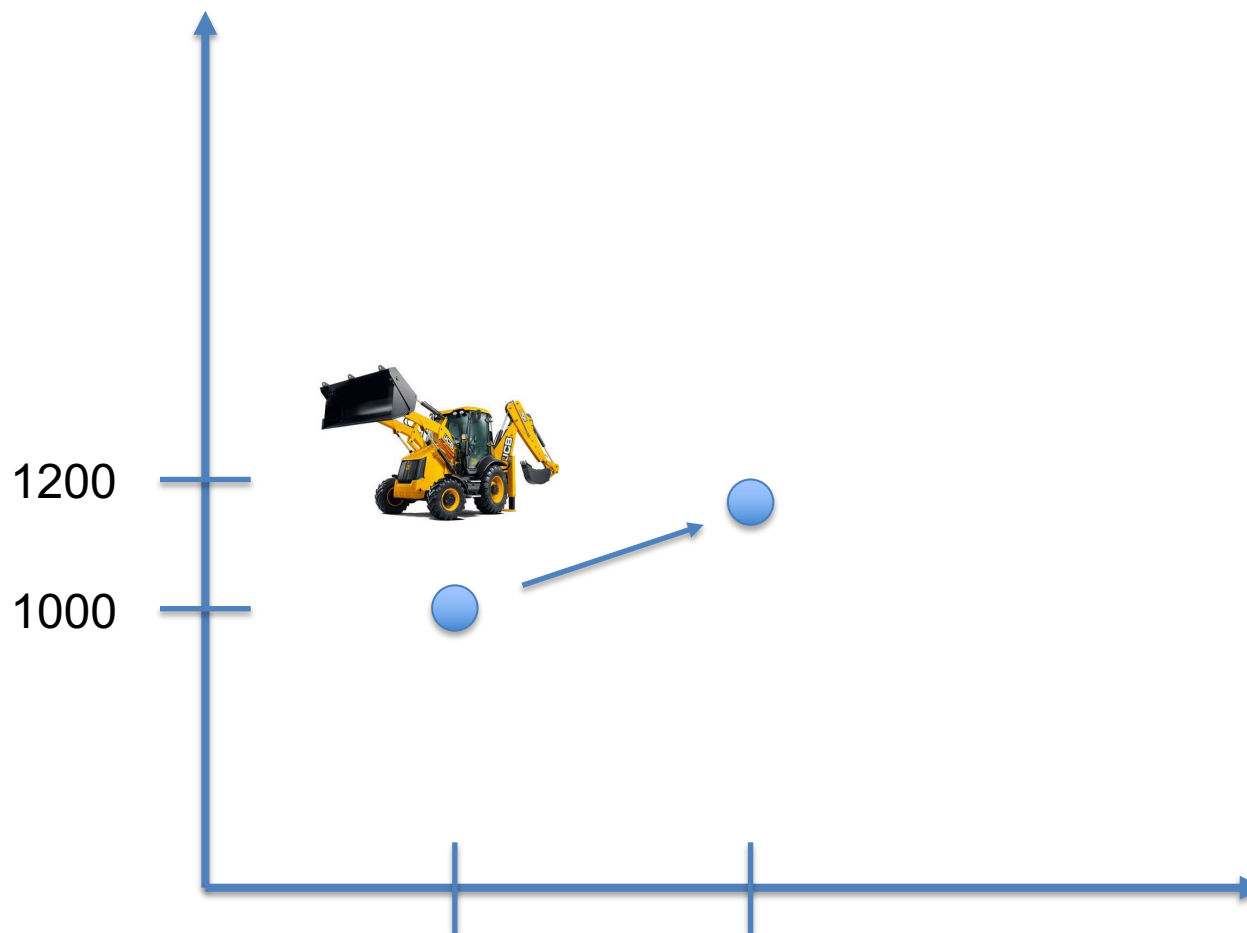
А если предыдущий месяц – май?



Хорошо или плохо?



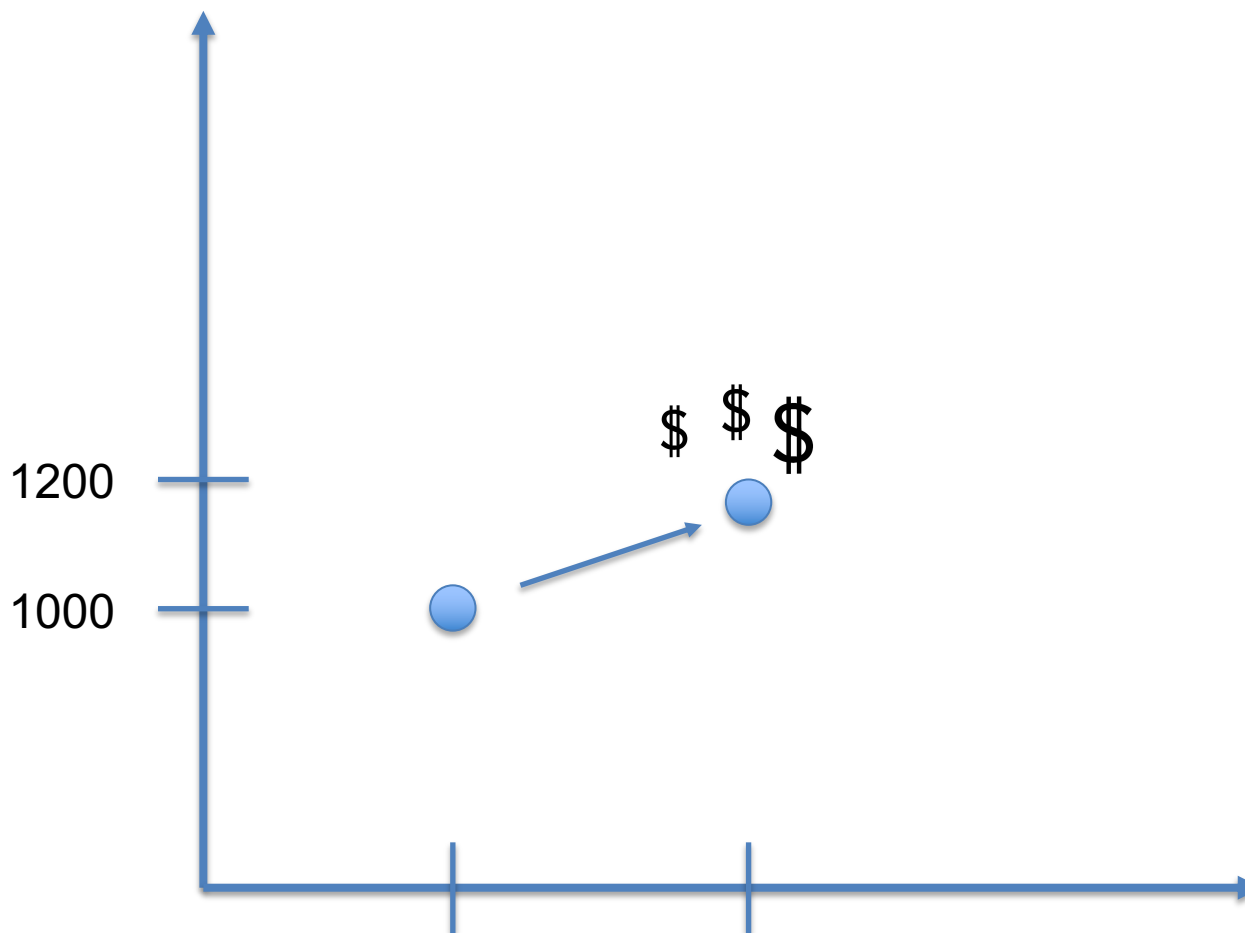
А если в предыдущем месяце вокруг булочной были перекопаны все дороги?



Хорошо или плохо?



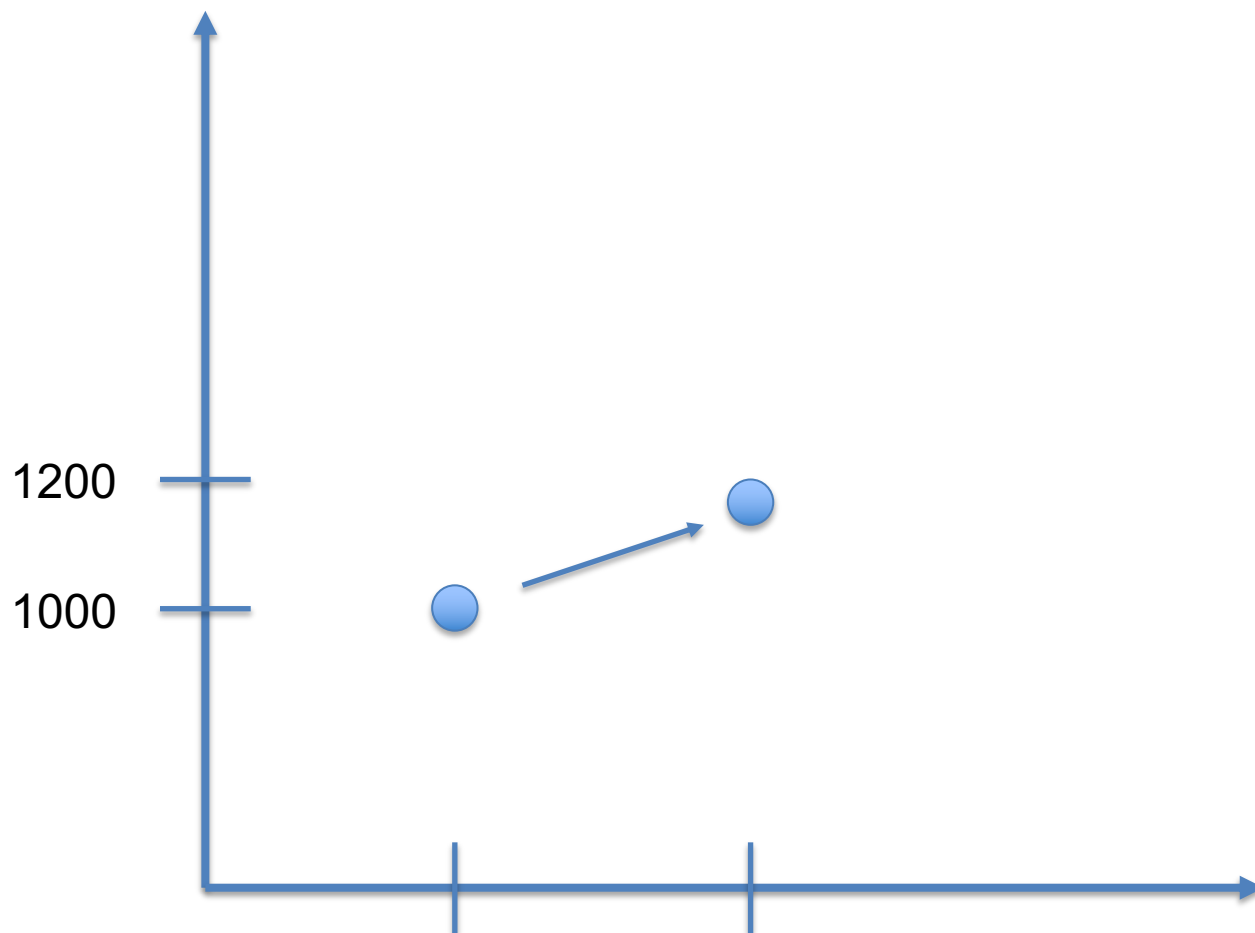
А если в этом месяце мы запустили супер дорогую рекламу, а в прошлом – рекламы не было совсем?



Хорошо или плохо?

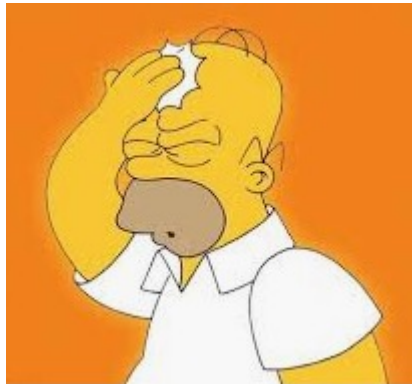


Что еще?



1. Твои ошибки (первая причина это ты)

- Смысловые: не разобрался что происходит в процессе, смотришь неправильные вещи
- Дизайнерские: плохо задизайнил тест, забыл залогировать рандом, неправильно выбрана целевая переменная.
- Технические: дубли, неучтенные миссинги, любые ошибки в коде





2. Качество данных

- хранилище не обновилось, поле не проставилось

client_id	client_age	min_sum	max_sum
31369064	99999	180	250
21505663	99999	270	480
17156604	99999	null	520
27452216	99999	210	540
58203623	99999	null	380
83225601	99999	null	560
34569524	99999	320	120

3. Календарные / сезонные

- выходные, праздники, НГ, короткий месяц февраль





Причины изменений

4. Единичные выбросы

- если средняя сумма покупок в одном из месяцев резко выросла, мб там один из клиентов купил самолет, а в остальном ничего не изменилось.
- как проверить: нарисовать распределение, посмотреть минимальное и максимальное значения в выборке.



5. Наблюдаемый поток изменился

- Меняется выборка: каналы, люди, клиенты, возраст...
- Другое распределение внутри потока
- Вспоминаем винтажный анализ и сегментацию



6. Внешние события

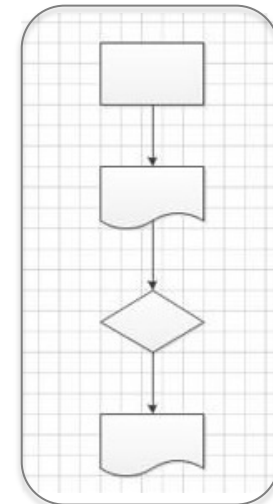
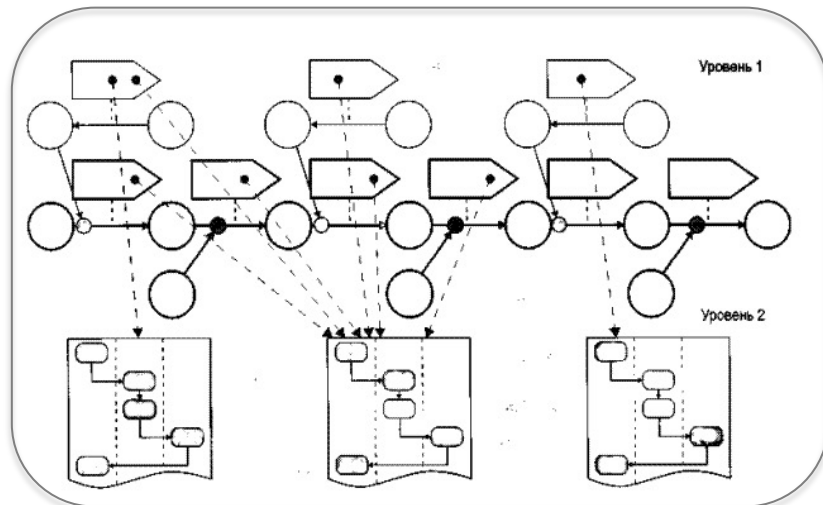
- Карантин, изменения законодательства, нововведения конкурентов, статьи в прессе...



Причины изменений

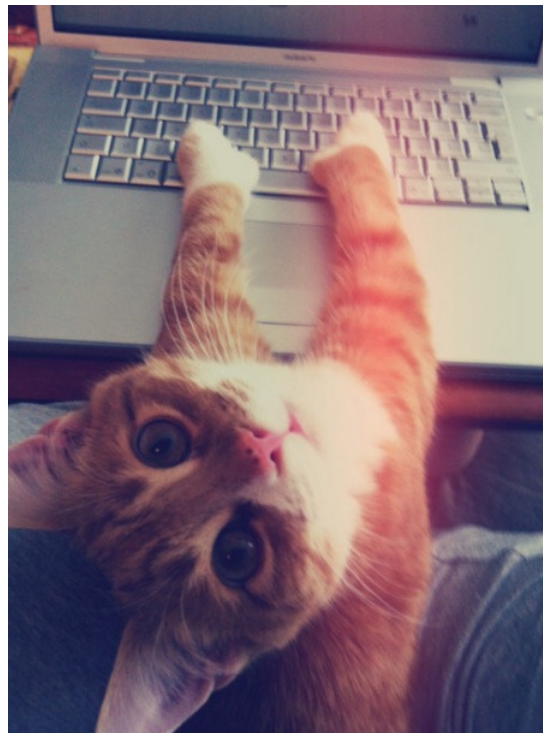
7. Внутренние изменения процесса

- Что-то стало работать по-другому
- Например, проводились тесты



8. Операционные

- Другие сотрудники обрабатывают звонки, или по другим инструкциям, или с другой мотивацией



9. Эффект дозревания





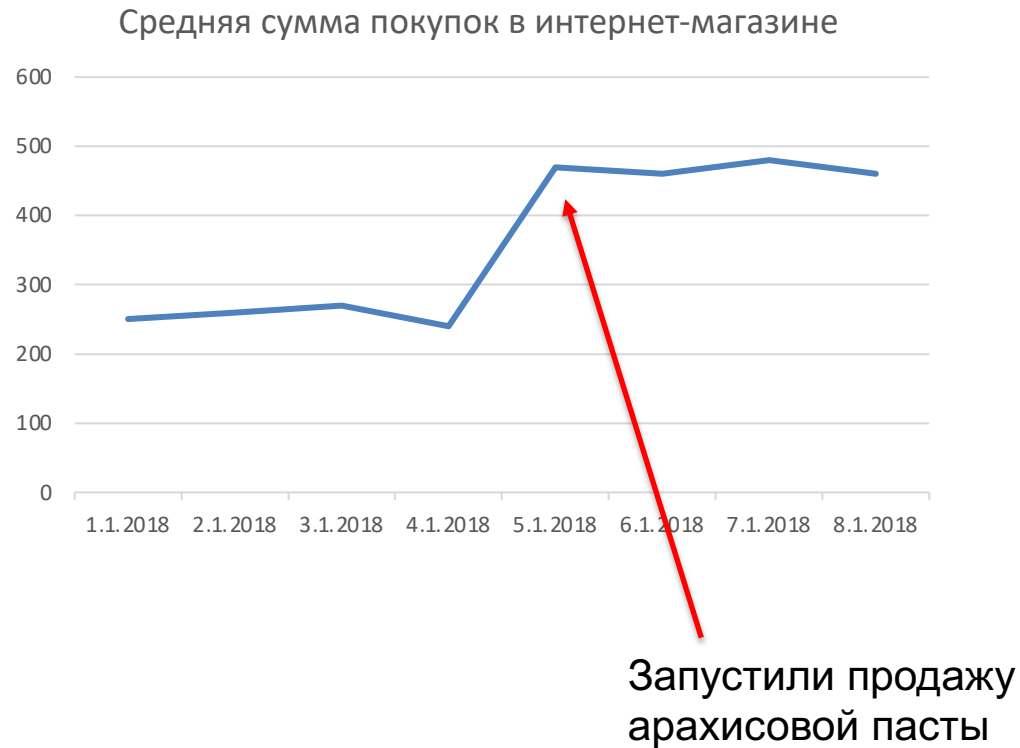
Что делать если видишь странный эффект?

- Проверь количество базы: достаточно ли объема выборки?
- Локализуй по времени. Изменение постепенное или одним днем?
- Сегментируй





«до-после» VS тест





Если не можешь принять решение по цифрам

- Всегда есть здравый смысл. Посмотри на процесс как пользователь или бизнесово.
- Выбери тот вариант, где нужно меньше доработок, процесс строится проще и риски будущих ошибок меньше.



Помни о цели





Правильно задавайте вопрос

? Сколько карт выдано?

- 1) Каких карт?
- 2) Где?
- 3) За какой период?
- 4) Зачем тебе эта информация? – мб нужен другой вопрос?



Как считать деньги: начало





PnL, NPV, LTV и другие непонятные сокращения

PnL

NPV

LTV

Юнит-экономика

Есть доходы, есть расходы. И конверсии.

Это не ракет саенс, не бойтесь, просто берите и считайте 😊



PnL подход

Profit and Loss Statement

- Бухгалтерский подход
- Формат отчетности
- По календарным периодам

ЗАЧЕМ

Видеть финансовое состояние компании/продукта, фактический кешфлоу продукта во времени

ОБЫЧНО

Считается для всего продукта целиком (но можно и по клиентам)

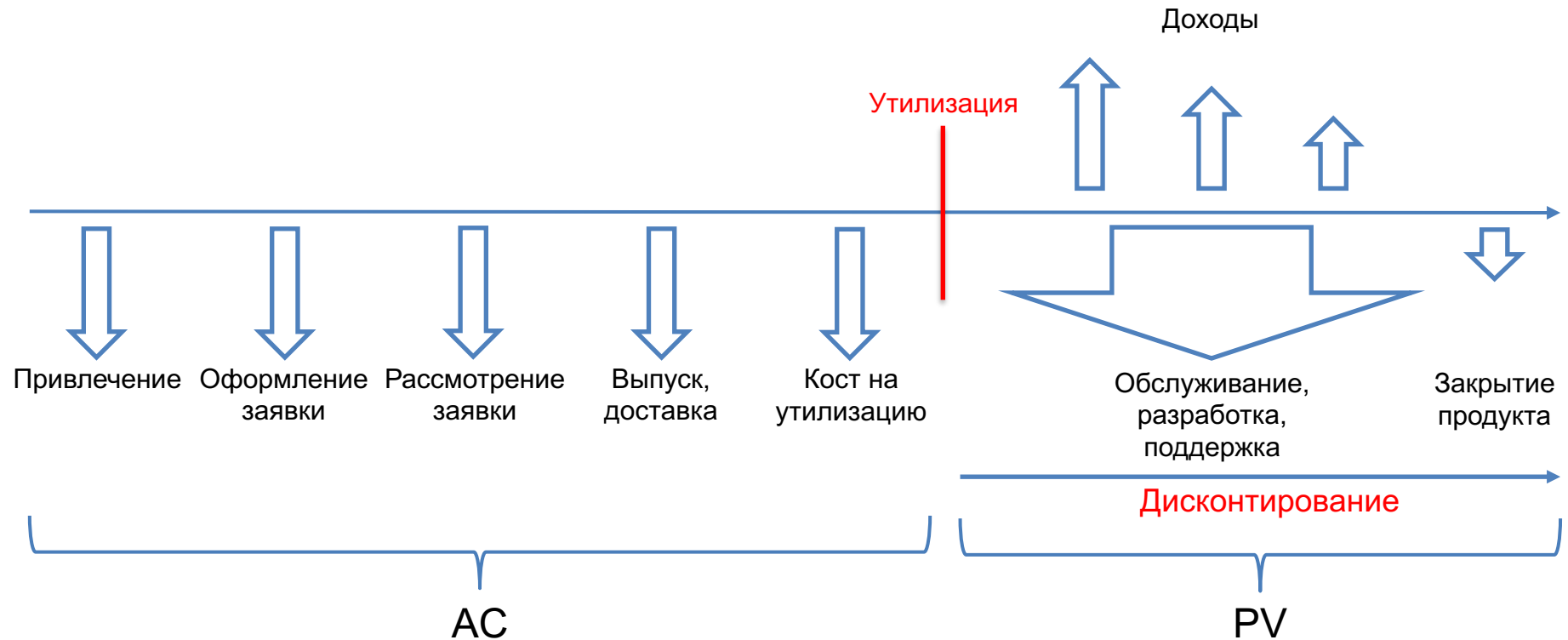
Product P&L				
Income Statement, mln rub				
Revenue				
Interest expenses				
Net interest and fee income				
LLP				
Net interest and fee income after provision				
Acquisition cost				
Transaction cost				
Service cost				
Operating income				
Extra cost				
Profit (loss) before def and oth				
Deferral effect				
Profit (loss) before FX and Tax				
Income tax				
FX (not related to products)				
Net income (loss) IFRS				

Временные периоды

Статьи доходов/расходов



$$NPV = PV - AC. \text{ Прогноз.}$$



$NPV > 0 \Rightarrow$ клиент прибыльный

ЗАЧЕМ

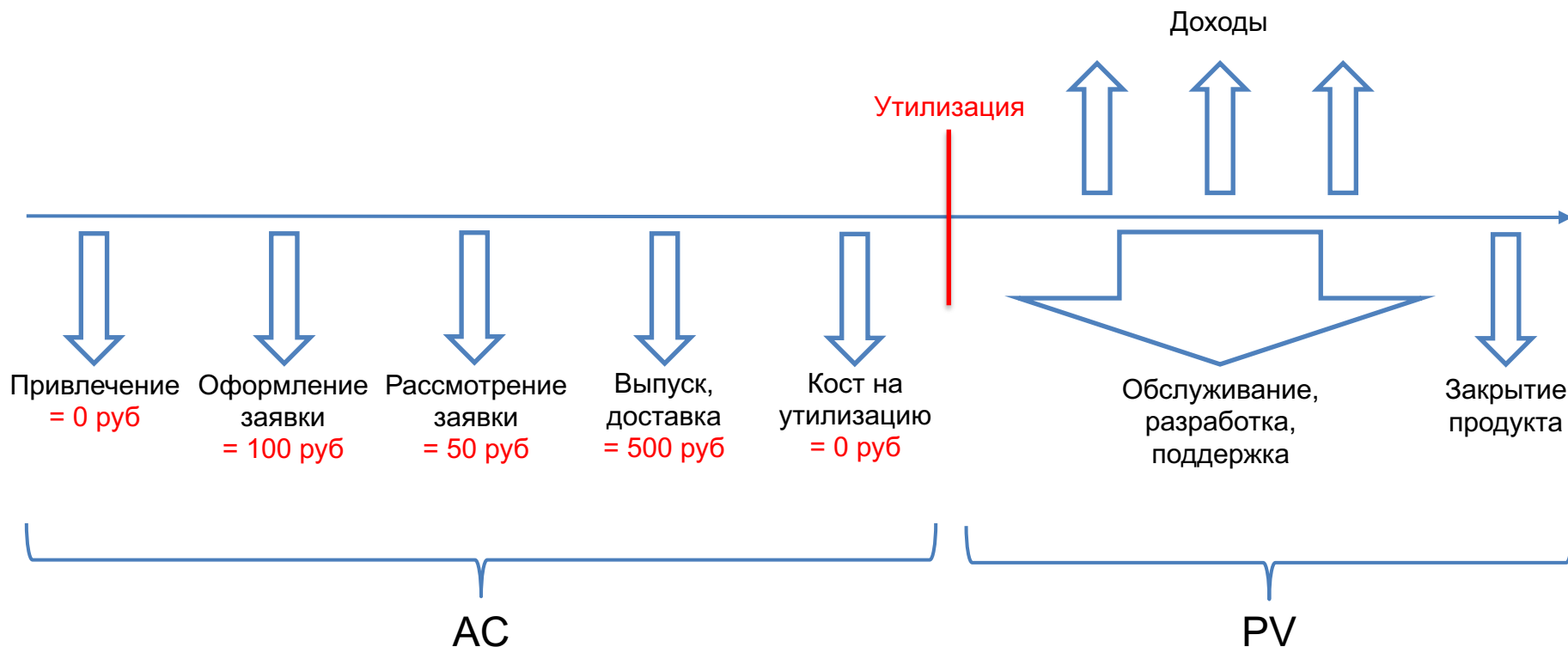
Принимать решения, прогнозируя прибыль/убытки

ОБЫЧНО

Считается для одного клиента (но можно и для всего продукта)

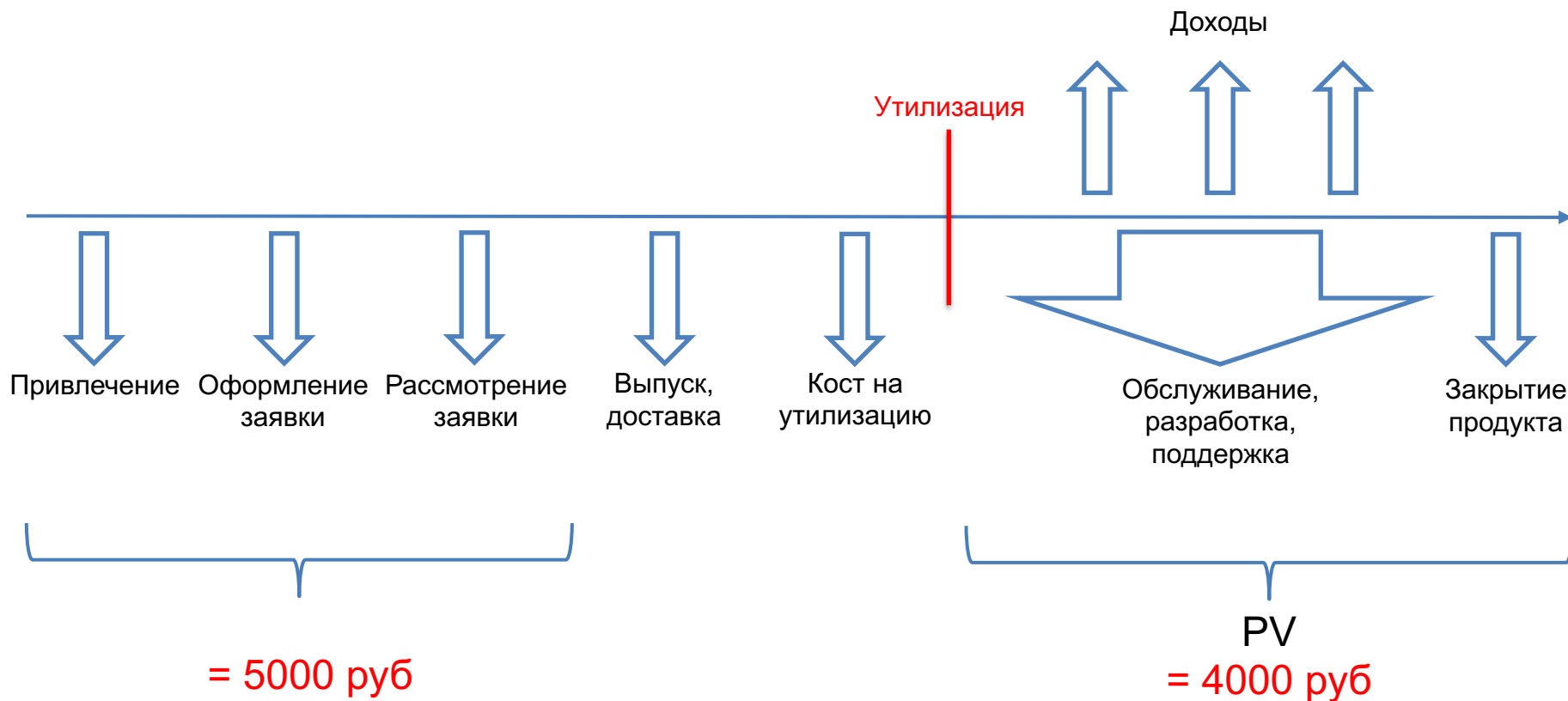


Пример 1



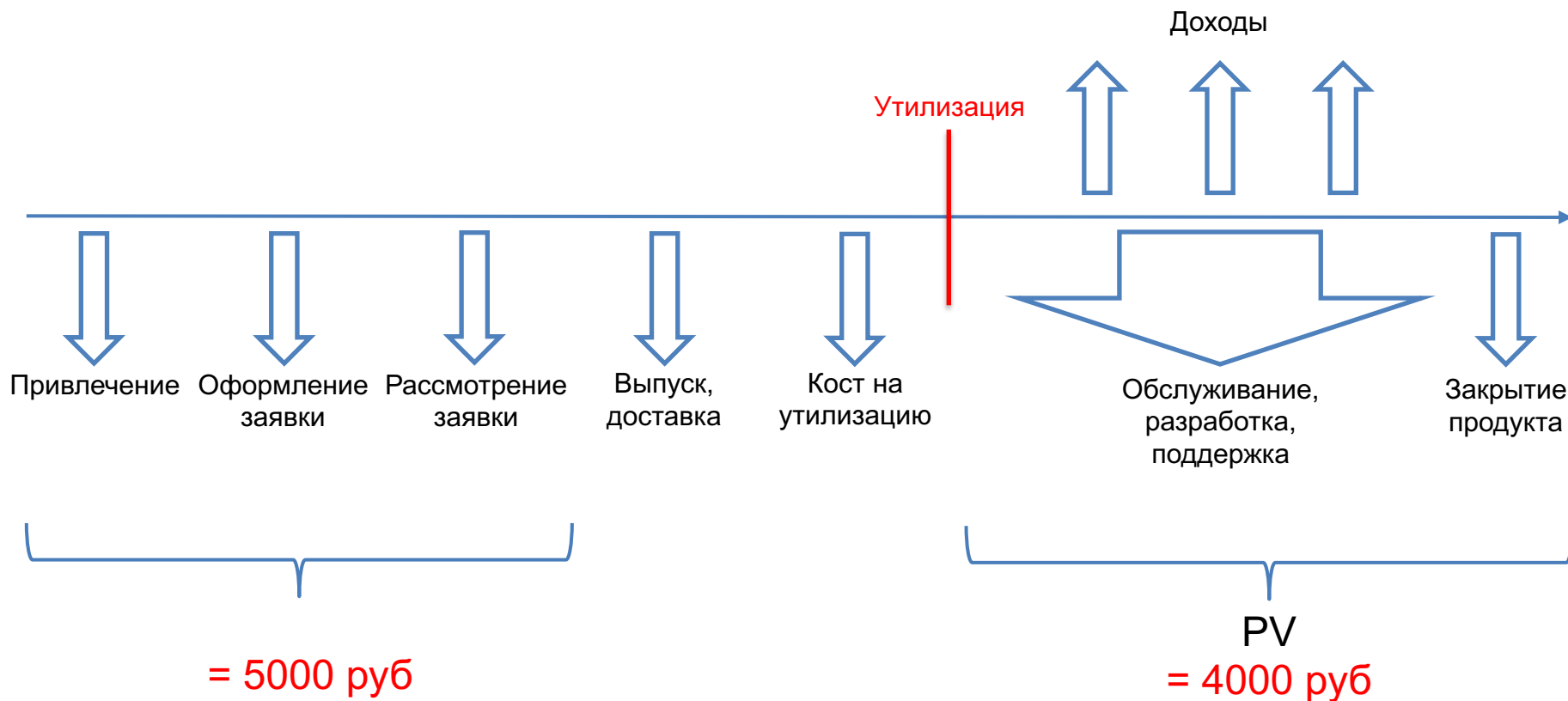
На схеме – стоимость прохождения одним человеком одного этапа. Клиент Иван прошел все эти шаги и только их, и утилизировался. Сколько стоит его привлечение?

Пример 2



На привлечение Ивана потратили 5К руб.
Оценили PV: заработаем с него 4К.
Нужно ли его одобрить?

Пример 2



PV – принятие решения по клиенту
NPV – принятие решения по каналу

Много аббревиатур, суть та же самая



Спасибо за внимание!

