

Анализ кейса «Сбермаркет»

Скрипт python3 для расчетов:

<https://clck.ru/UqZw7>

Дизайн эксперимента

Ошибки дизайна эксперимента 1

Нет «бизнес проблемы»

- Из постановки задачи нет ясности «бизнес проблемы». Есть только «идея». Как следствие, нет уверенности, что выбранные метрики, действительно, описывают что-то важное для бизнеса. Возможно, нужны другие метрики.

Примеры потенциальных проблем и метрик:

Пример1

«x% **пользователей** от начавших оформления заказа уходят с сайта из-за того, что они сталкиваются с модалкой выбора адреса после добавления товара в корзину. Если мы изменим сценарий, то эта доля сократится до у%»

Конверсия из Main Page в добавление товаров **по уникам**. В этом примере мы **теряем клиентов**, которые не могут получить доставку.

Пример2

«x% **сессий**, в которых начали оформлять заказ не заканчиваются добавлением в корзину из-за модалки. Если мы изменим сценарий, то доля сократится до у%».

Конверсия из Main Page в добавление товаров **по сессиям**. В этом примере мы **теряем заказы**, потому что пользователям слишком сложно завершить заказ.

Пример3

«Мы теряем повторных покупателей из-за потенциально нелогичного сценария. Если мы изменим сценарий, то повторные покупки вырастут с x% до у%».

Retention 2-й покупки.

Ошибки дизайна эксперимента 2

Два изменения в одном эксперименте вместо одного

- В эксперименте было внесено сразу 2 изменения:
 - Landing Page (другой оффер и дизайн)
 - Последовательность шагов (целевой тест)

Даже при сохранении отказов в LP тестовые пользователи получают другой оффер, который может повлиять на ожидания пользователей, а значит, на тестируемую гипотезу.

После полученных результатов останутся вопросы, что повлияло LP, или последовательность шагов.

Ошибки дизайна эксперимента За На кого нацелен эксперимент?

- В эксперименте не указано на каких пользователей нацелены изменения. Примеры разных типов пользователей:
 - делали уже покупки - не делали
 - с какого типа устройства
 - ранее были на сайте – или впервые пришли
 - с какого канала РК были пользователи, из какого георегиона и т.д.

Пример: если человек ранее делал покупки и знает, что его адрес есть в списке доставок, то изменение flow будет влиять по-другому нежели на новичков или на тех пользователей, которые заведомо знают, что их адреса нет в списке доставки. Поведение с мобильника отличается от десктопа и т.д.

Совет для новичков аналитиков: мы должны точно понимать на какой пользовательский опыт мы действуем и как, а не на всех пользователей «вообще»

Ошибки дизайна эксперимента 3б

На кого нацелен эксперимент?

- Адрес доставки – атрибут описывающий профиль пользователя. Если у человека больше несколько 10-и адресов – это аномальность, а скорее их менее 5 (пока оценочно).
- С другой стороны возможность доставки на адрес - это атрибут профиля магазина.
- Т.о. планирование эксперимента, а именно конверсии в добавление в корзину не может быть без привязки к магазинам

Важный вывод:

- Если у магазина доля ареала доставки от ареала желаемых адресов пользователей большая, то метрика конверсия описывает удобство пользования сайтов
- Если у магазина эта доля мала, то помимо удобства пользования сайтом появляется (не)возможность доставки. В одном случае человек должен отказаться от собранных товаров, в тестовом он может поменять магазин до выбора товара. В задаче не ясно, а какую проблему мы решаем?

Ошибки дизайна эксперимента 4

Нет критериев успеха и стартовых метрик

- «Кривой» сценарий оценочное суждение 😊
- Нужно понимать стартовые метрики еще при планировании эксперимента, условия их получения.
- Нет понимания какой результат будет приемлем (какое увеличение «конверсии в добавление корзину» необходимо принять, чтобы это было коммерчески выгодным).

Совет новичкам аналитикам: неопределенный на старте приемлемый результат влечет к отсутствию расчетов необходимых размеров выборок, а значит, потенциально к «эффекту подглядывания», а также ошибок 2-го рода (принять неверную нулевую гипотезу об отсутствии различий результата)

С точки зрения бизнеса также нет понимания насколько выгодно будет раскапывать в прод тест в случае успеха. Выкатить в прод стоит денег... и если ожидаемое увеличение конверсии, например, стат. значимо +0,5%, то имеет ли выкатывать в прод смысл?

Ошибки дизайна эксперимента 4

Технические

Сразу бросается в глаза, что отсутствуют идентификатор сессии. Это значит, что мы можем лишь довериться, что фактических перескоков из одной группы в другую во время взаимодействия пользователей с сайтом не было. В идеале иметь справочник: Сессия – Факт показа тест/контроля сайта.

Дизайн экспериментов

Здесь 2 эксперимента, поэтому необходимо для каждого разработать дизайн.

- Изменение Лендинга
- Изменение последовательности выбора адреса

Дизайн эксперимента LP1

Гипотеза

Принципиальное изменение LP не приведет к статистически значимым ухудшениям по отказам новыми пользователями с рекламного канала X из региона Y

... и не повлияет на добавление товары и оформления заказов*

Метрики

Отказы

конверсия в добавление товара*

конверсия в оформление заказа*

Сравнить конверсии на разных типах устройств, браузерах

**формулировки, которую имело бы смысл добавить до факта эксперимента*

Дизайн эксперимента LP1

Методология

Меняем только LP, все оставляем прежним. Работаем только с новыми пользователями из одного региона с одного рекламного канала.

Минимальные условия принятия

конверсия в оформление заказа не ухудшится

конверсия в добавление товара не ухудшится

изменение отказов не ухудшатся более +1 п.п. с дов.вер 95% и
мощн.80%

Дизайн эксперимента LP3

Контроль

СБЕР МАРКЕТ

+7 800 222 11 00 Войти

Доставим продукты и товары из магазинов

Выберите магазин и начинайте делать покупки

METRO **ЛЕНТА** **АЗБУКА ВКУСА** **Ашан** **ВкусВилл**

Аллея **Бахетле** **ЗЕЛЕНОЕ ЯБЛОКО** **МЕГАМАРТ** **ВИКТОРИЯ**

командор **real**

Забудьте про походы в магазины

Мы сохраним ваше время, энергию и деньги для более... Покупайте по скидкам и акциям... Соберёшь лучшее с полок магазина, Доставим день в день, в удобное время... Донесём тяжёлые сумки прямо...

Тест

СБЕР МАРКЕТ

+7 800 222 11 00 Войти

Доставим продукты и товары из магазинов

Введите адрес доставки и мы покажем, какие магазины вам доступны

Город, улица, дом

Партнёры

METRO **ЛЕНТА** **АЗБУКА ВКУСА** **Ашан** **ВкусВилл**

Аллея **Бахетле** **ЗЕЛЕНОЕ ЯБЛОКО** **МЕГАМАРТ**

Показать всех

Забудьте про походы в магазины

Мы сохраним ваше время, энергию и деньги для более важных вещей и событий... Покупайте по скидкам и акциям наших партнёров... Соберёшь лучшее с полок магазинов, как для себя... Доставим день в день, в удобное время... Донесём тяжёлые сумки прямо до двери...

Скрипт python3 для расчетов: <https://clck.ru/UqZw7>

Дизайн экспер-та последовательность1

Проблема

Мы теряем заказы новых пользователей из региона Y с канала X, которые тратят свои силы на добавление заказа, а им не возможно доставить товар.

Гипотеза

Изменение последовательности шагов для новых пользователей из региона Y с канала X позволит пользователям осознанно выбрать магазин с доставкой из их региона (в случае отсутствие желаемого переключиться на другого ретэйлера прежде, чем юзер потратит время на поиск товара).

Метрики

Конверсия в добавление товара из LP

Конверсия в оформление товара из LP

Сравнить конверсии на разных типах устройств, браузерах, осей

Дизайн экспер-та последовательность1

Методология

Меняем только последовательность шагов, все оставляем прежним. Работаем только с новыми пользователями из одного региона с одного рекламного канала.

Расчитываем размер выборки для 95% дов.вер и 80% мощности эксперимента из мин. условий принятия решения.

Ставим трекер сессий, чтобы отслеживать пользовательский опыт.

Минимальные условия принятия

конверсия в оформление заказа улучшится на +Ап.п. (процентные пункты)

конверсия в добавление товара улучшится на +Вп.п.

В этом случае мы получим +Z р. за следующие N мес., что отобьет инвестиции в исследование и разработку.

Проверка данных

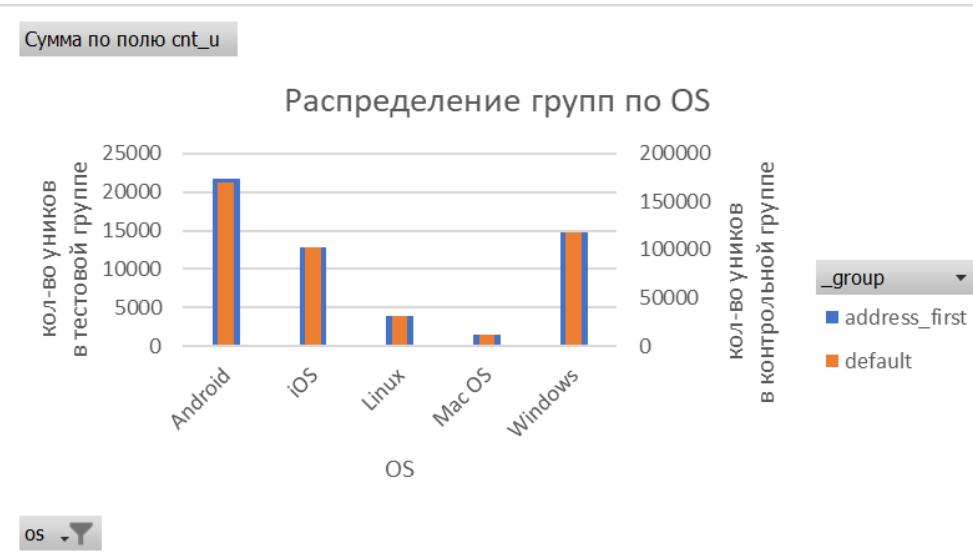
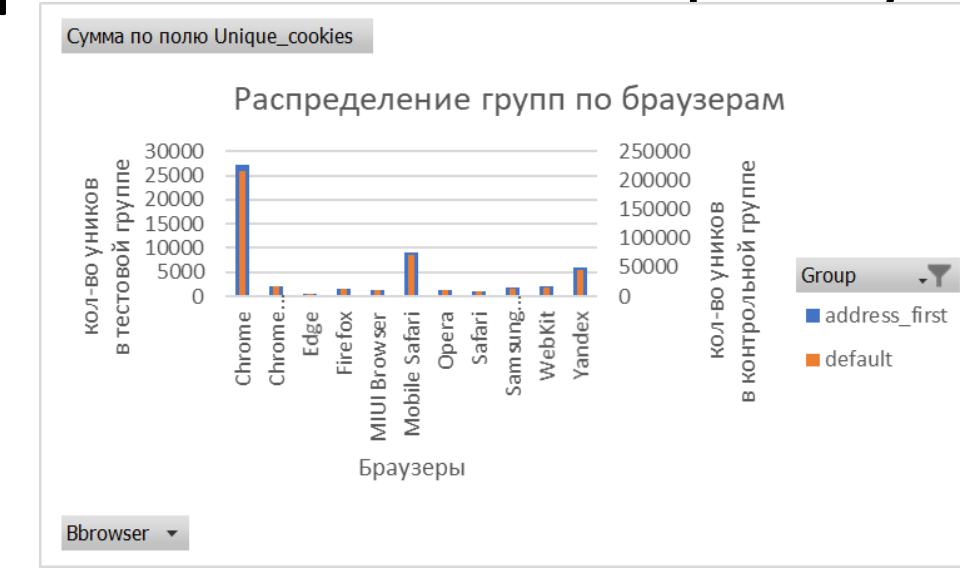
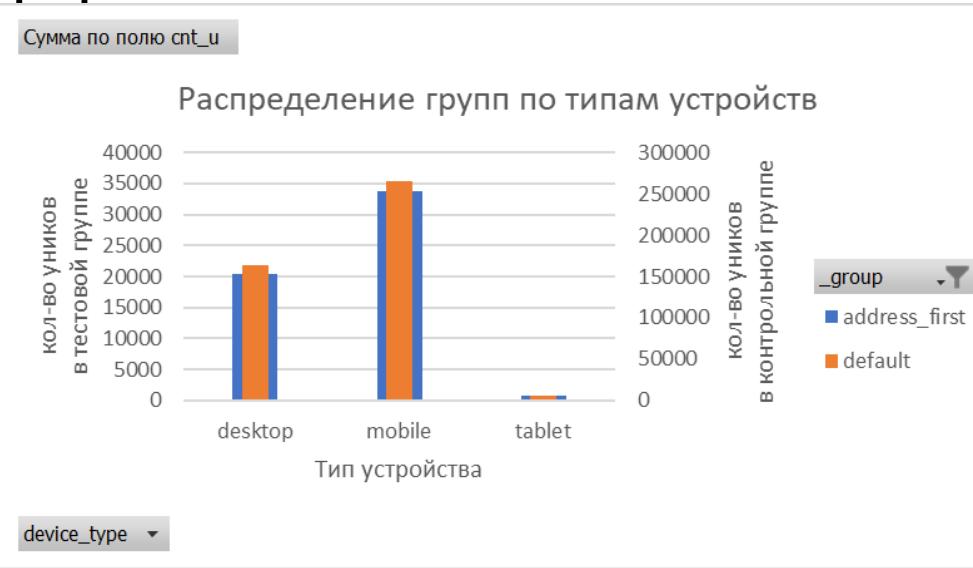
Деление на тест и контроль

Проведенный анализ показал, что деление на тест и контроль был выполнен сбалансировано. Доли типов устройств, браузеров, OS в обоих группах сохраняется

Но с учетом предыдущих ошибок дизайна это лишь констатация приятного факта, нежели полезный вклад

На следующем слайде приведена визуализация этого факта + скрипт в PostgreSQL

Деление на тест и контроль по тех.атрибутам



```

select *,  

    sum(cnt_u) over (partition by _group) qnt_gr,  

    cnt_u/sum(cnt_u) over (partition by _group) perc_  

from  

(  

    select _group  

        ,device_type  

        ,browser  

        ,os  

        ,count(anonymous_id) cnt_all, count(distinct anonymous_id) cnt_u  

    from ab_test_hit  

    group by _group  

        ,device_type  

        ,browser  

        ,os  

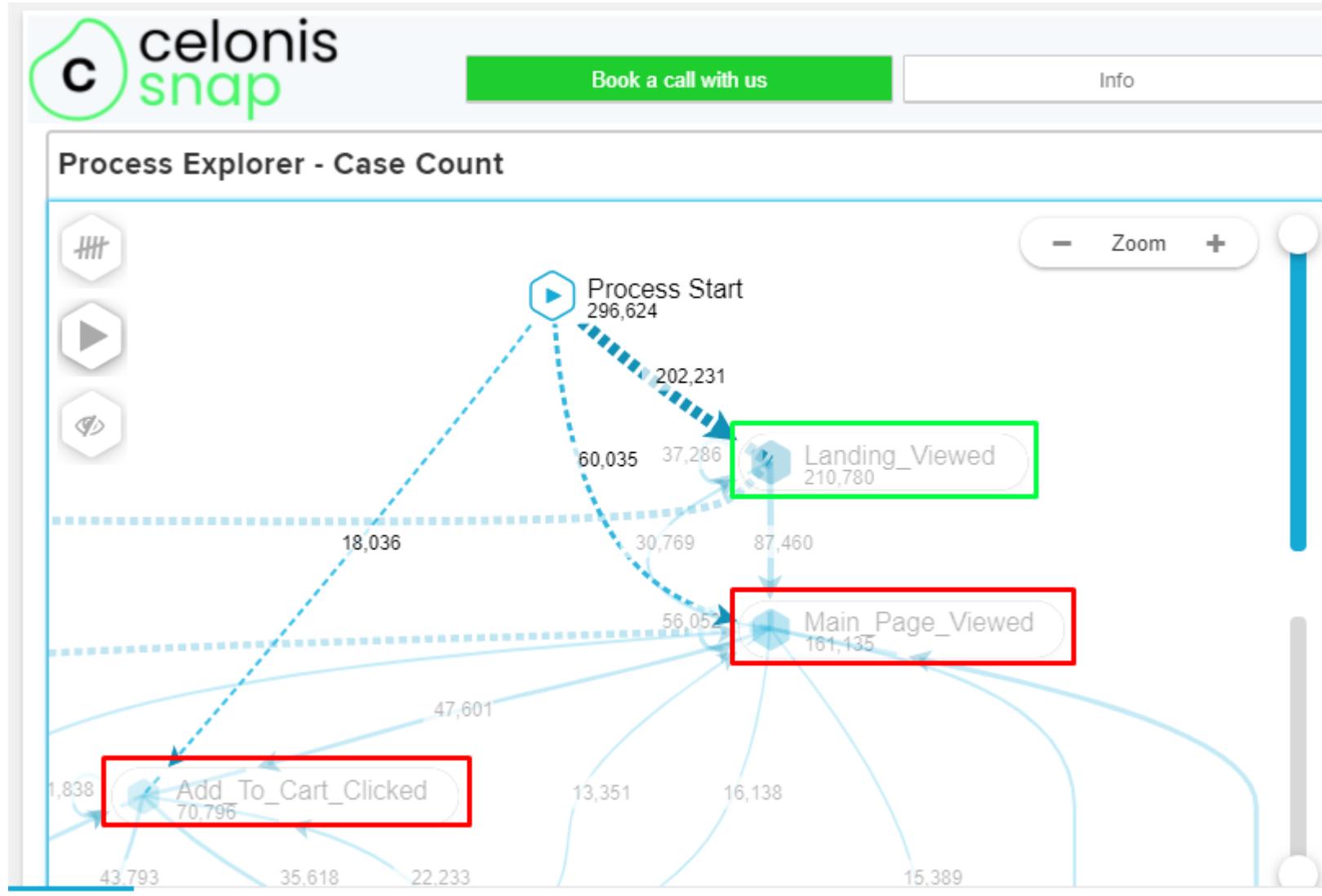
) f  

order by device_type,browser,os,_group;
  
```

Граф последовательности событий

1. Собрали все Events в python
 1. timestamp
 2. anonymous_id
 3. Event = ‘одноименные названия таблиц’
2. Отправили в приложение для графа Celonis
3. На следующих слайдах смотрим на последовательности в графе и делаем выводы

Граф последовательности событий



32% сессий были начаты не с Landing Page

Для анализа возьмем только тех пользователей, которые начали свой flow с LP (считаем их новыми)

Оставим только последовательности, в которых первый шаг LP, второй либо открытие MainPage, либо модалка, либо конец сессии

*celonis дает качественное представление ситуации

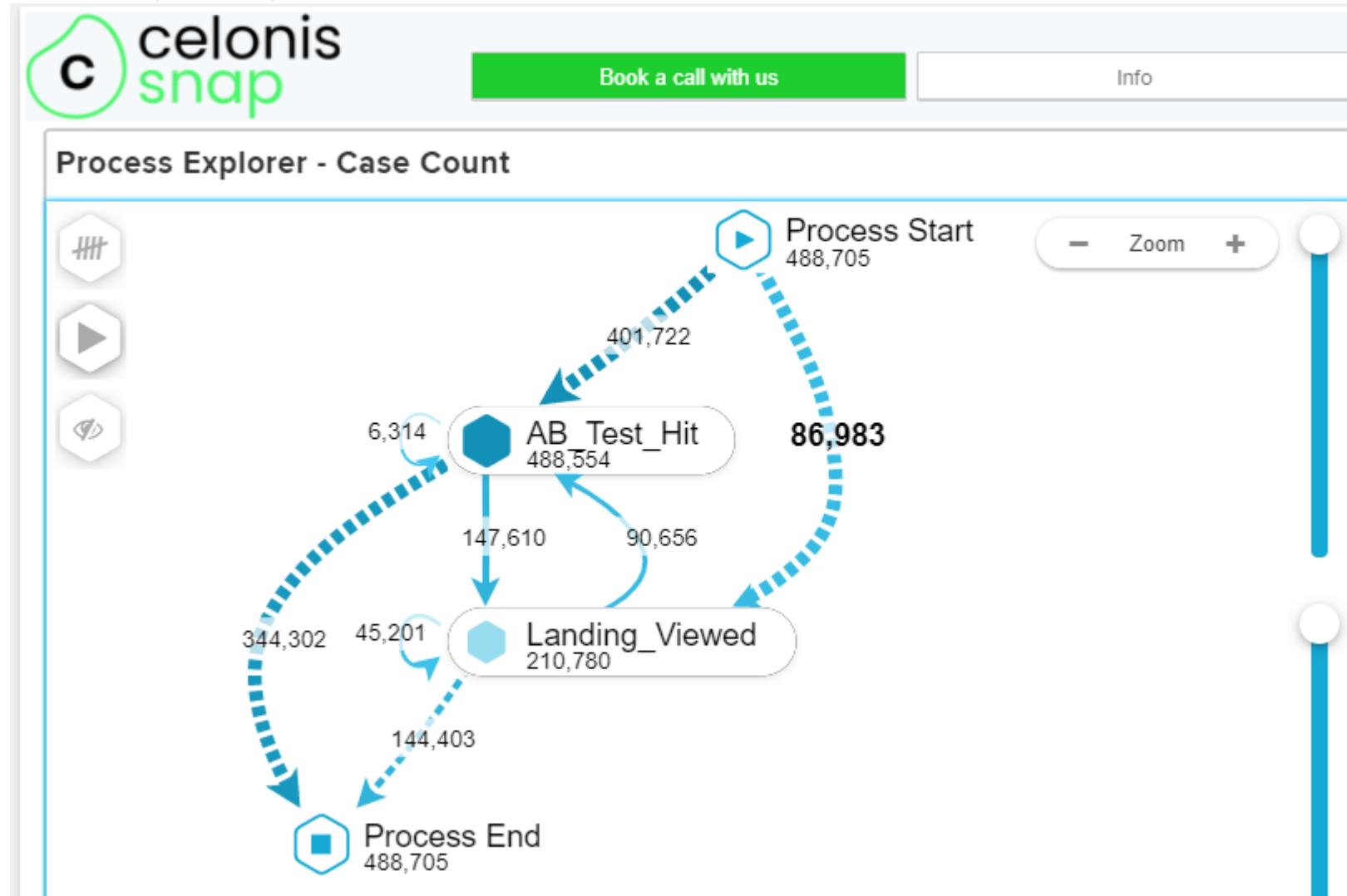
Скрипт для исключения Ид

Аномальные anonymous_id сохранены в exclude1

```
1 temp = events.sort_values('timestamp').groupby('anonymous_id',as_index=False).agg({'Event':lambda x: list(x)})
2 temp['step1'] = [x[0] for x in temp['Event']]
3 temp['step2'] = [x[1] if len(x)>1 else 'End' for x in temp['Event'] ]
4
5 good = set(temp.query('''(step1 == 'Landing_Viewed') & ((step2=='Landing_Viewed')|(step2=='End')|(step2=='Main_Page_Viewed')
6 anonymous_id)
7 mask = [x not in good for x in temp.anonymous_id ]
8 exclude1 = set(temp.anonymous_id) - set(good)
```

Для каждой дальнейшей аномалии будет формироваться set с именем excludeN
Далее они объединяются в один set и исключаются из все выборки

Граф последовательности событий



Деление на тест и контроль произошел:

- 401k пользователей до LP
- 86k пользователей после LP

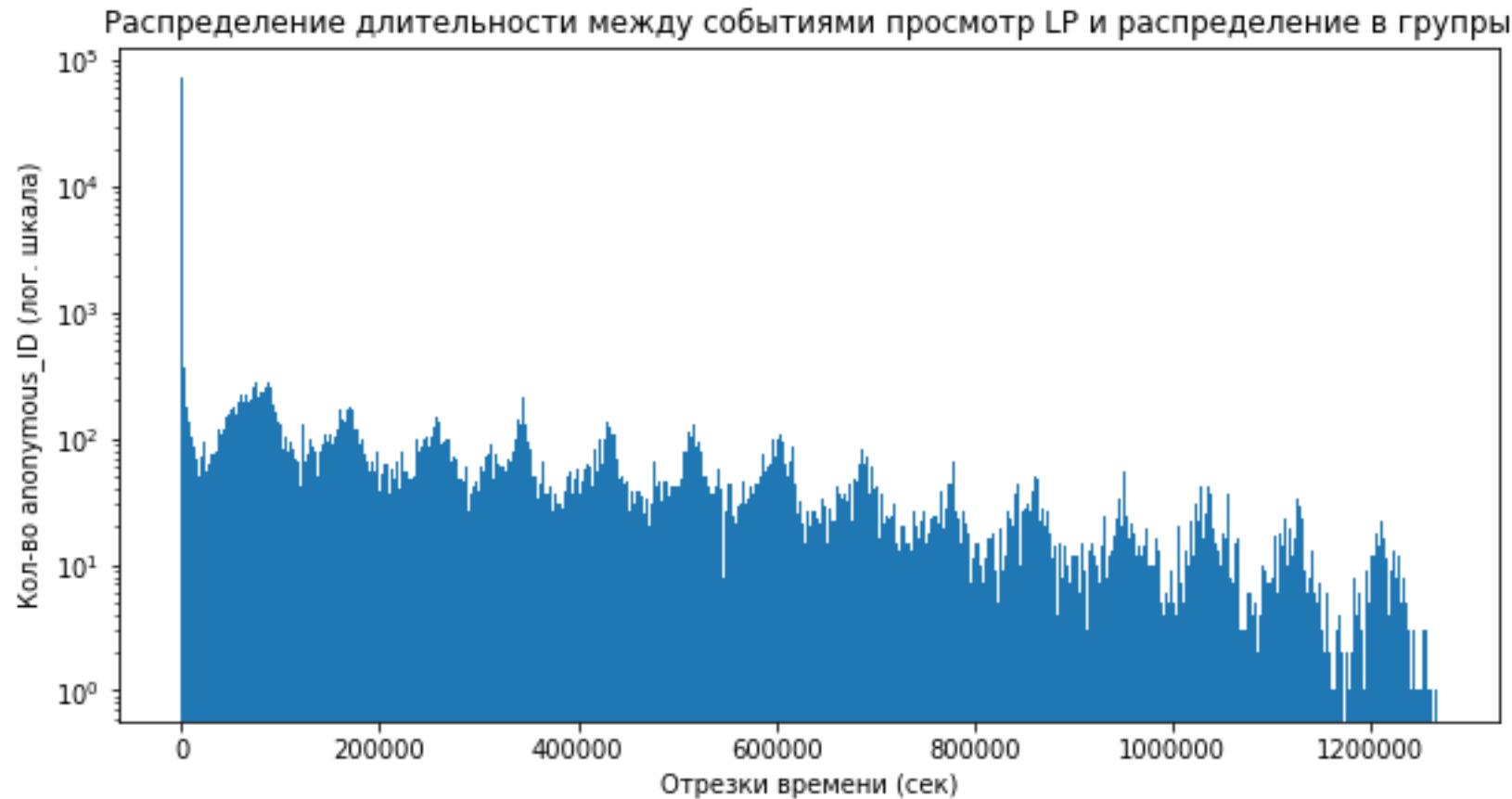
Потенциально из-за задержек события могут прилетать не последовательно Поэтому Δсек по модулю ~5 сек нормально

Далее изучим этот вопрос и возьмем только те, в которых время события $tstamp(LP) \geq (tstamp(AB_test_hit) - \Delta\text{сек})$

Δсек изучим отдельно

При $\Delta\text{сек} > 5$ сек ситуация аномальна

Просмотр LP до AB_test_hits



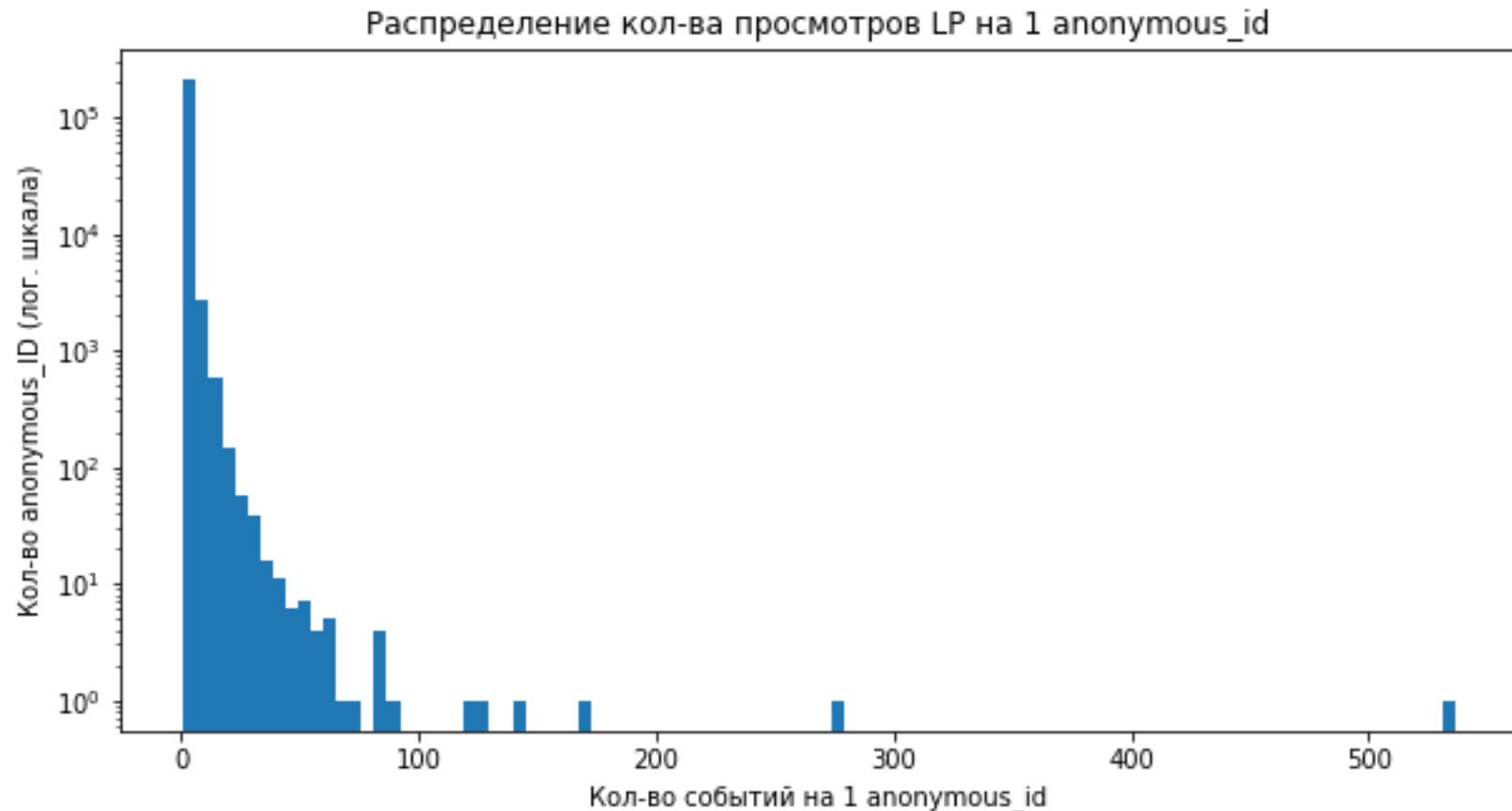
Алгоритм построения

1. Anonymous_ID, в которых события LP раньше AB_test_hits
- 2 Построение гистограммы распределения
- 3 Ось Y логарифмическая

Вывод: необходимо удалить ID, в которых Δsek > 5 сек

5 сек выбрано эвристически

Landing_Viewed



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 8

Будем считать кол-во событий на один anonymous_id >8 аномальными

Удалим аномальные anonymous_id

Резкое падение событий LP



Резкое падение событий LP после
10.12.20

Гипотеза ситуации:
Завершение платного трафика на LP

Address_Change_Initiated



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 8

Будем считать кол-во событий на один anonymous_id > 8 аномальными

Удалим аномальные anonymous_id

Address_Change_Initiated

source	count	unique	rate
add_alcohol	283	218	77%
add_product	57265	42326	73%
address	31360	15574	49%
landing	13083	8763	67%
not_in_delivery_zone	1443	1118	77%

	source	group	unique
add_alcohol	address_first	33	
	default	185	
add_product	address_first	3299	
	default	39054	
address	address_first	1358	
	default	14232	
landing	address_first	8741	
	default	47	
not_in_delivery_zone	address_first	90	
	default	1028	

Построена сводная таблица распределения неуникальных и уникальных событий по разным источникам.

Вывод1. В тестовой и контрольной группе попадают события, которые там не должны быть. **Исключим**, т.к. если у an...s_id был пользовательский опыт из разных групп, то его опыт использовать нельзя.

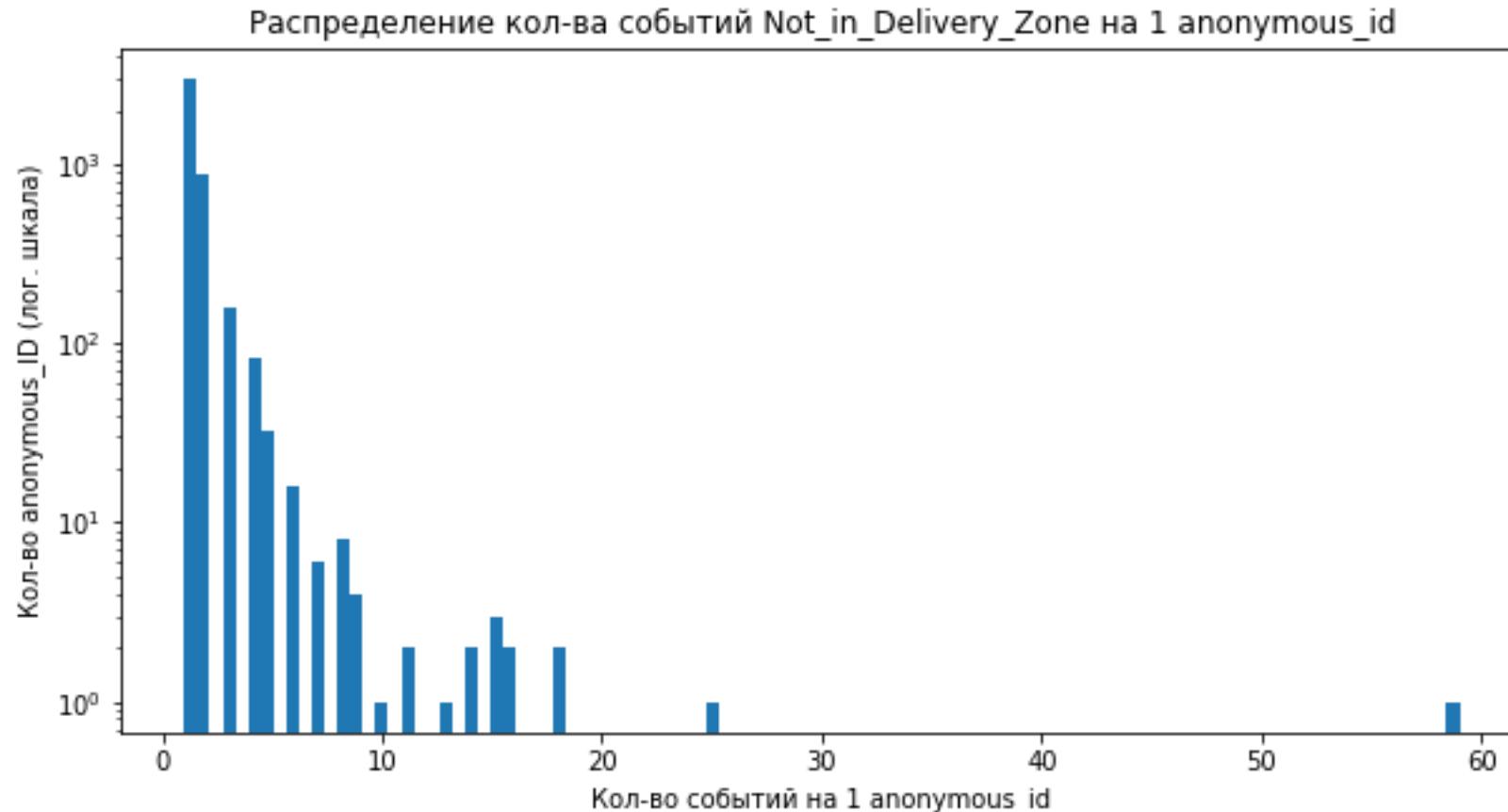
Объясняться может тем, что an...s_id уже заходил на сайт до того, как попал в группу. Этот момент не проверялся на данном этапе.

Вывод2. Есть события, которые не описаны в задаче: «address», «add_alcohol». Возможно, это какие-то стандартные хиты для добавления в корзину товара. Тем не менее, **пока данная информация не прояснена у оунера**, предлагается их исключить.

Вывод3: Важно иметь session_id со справочником к какому эксперименту сейчас относятся, т.к. иначе нет ясности в других таблицах, к какому эксперименту данных хит, с учетом потенциальных ошибок и накладок...

Скрипт python3 для расчетов: <https://clck.ru/UqZw7>

Address_Not_in_Delivery_Zone



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

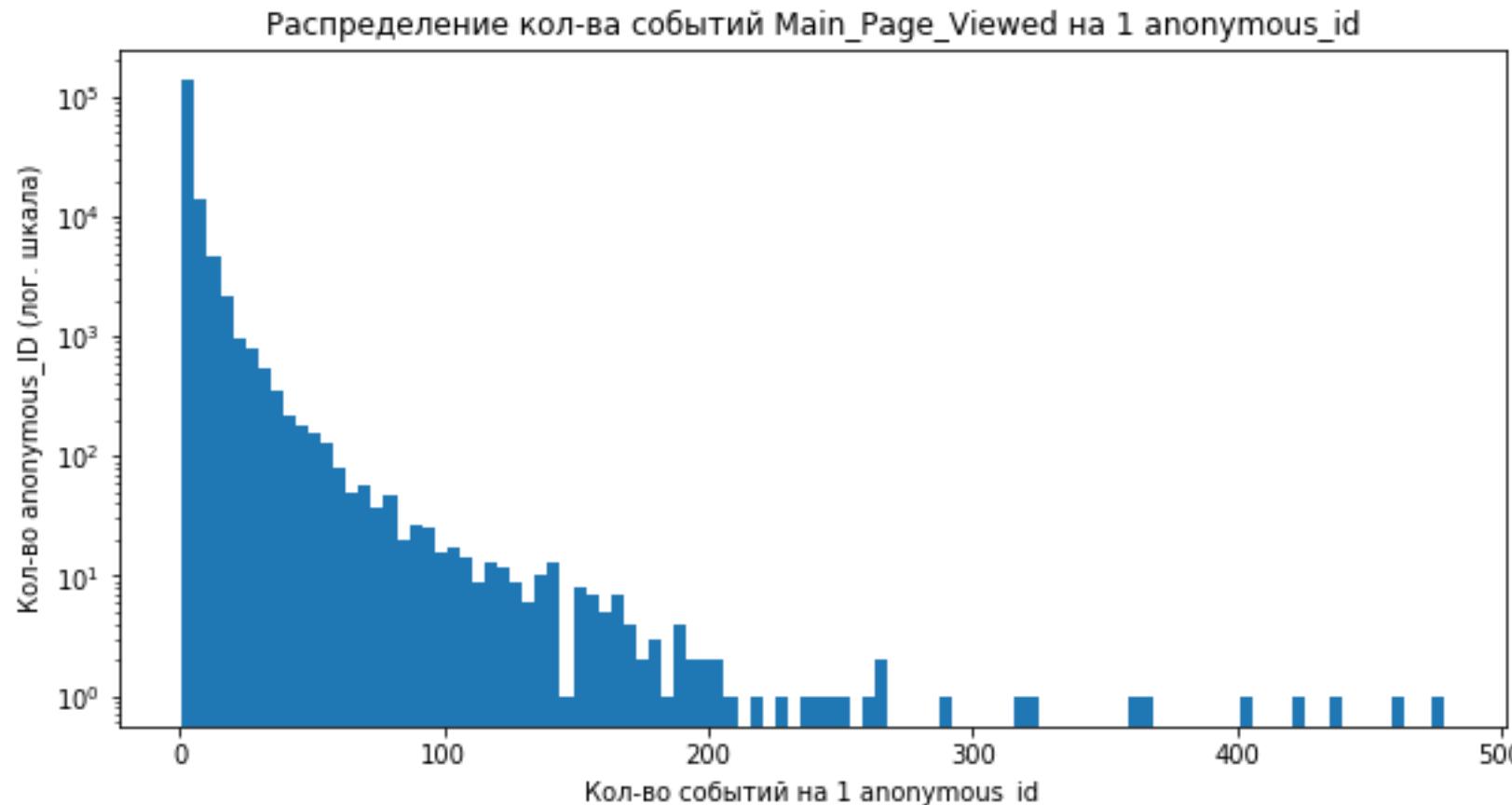
Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 6

Будем считать кол-во событий на один anonymous_id >6 аномальными

Удалим аномальные anonymous_id

Main_Page_Viewed



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 34

Будем считать кол-во событий на один anonymous_id > 34 аномальными

Удалим аномальные anonymous_id

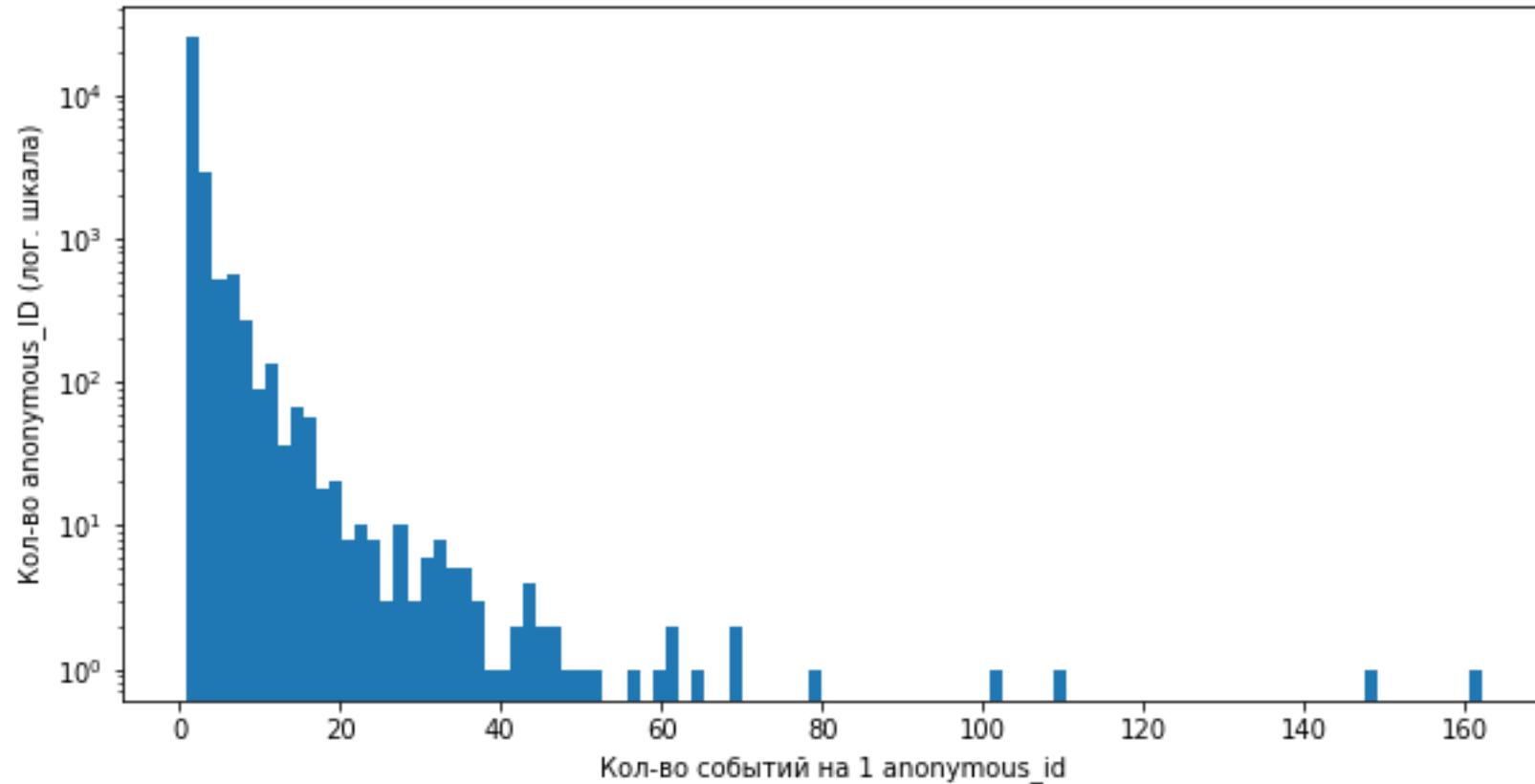
Main_Page_Viewed

retailer_id	count
1	218977
29	150384
57	108035
93	19502
33	12031
97	10287
25	6855
193	5750
69	5342
229	5140

Обратим внимание, что магазины 1,29,57 явные лидеры по числу событий

Shop_Celection_Started*

Распределение кол-ва событий Shop_Celection_Started на 1 anonymous_id



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 12

Будем считать кол-во событий на один anonymous_id > 12 аномальными

Удалим аномальные anonymous_id

* Грамматическая ошибка в "Celection" сохранена, как в источнике

Shop_Selection_Started

source	count
address_modal	14290
header	36553
landing	6564

source	group	Count_in_group	Count_total	rate
address_modal	address_first	1196	15975	0.07
	default	14779		0.92
header	address_first	27828	252203	0.11
	default	224375		0.88
landing	address_first	8024	8060	0.99
	default	36		0.004

Построена сводная таблица распределения неуникальных событий по разным источникам и по группам для теста.

Вывод1. Есть ошибки деления на группы. Тестовая группа попадала на путь контрольной и наоборот. Исключим.

Вывод2. Похоже, проводился еще эксперимент параллельно (событий Header больше 50%). Пропорция внутри Header между тестом и контролем говорит о том, что его нельзя отнести ни к тесту, ни к контролю. Лучше расчеты провести без этого источника, **исключив все anonymous_id с ним. Тем не менее необходимо прояснить это событие у Оунера.**

Если параметр указывает на независимый эксперимент, то необходимо сбалансировать присутствия *an...s_id* из этих экспериментов в teste и контроле, чтобы сохранилась пропорция.

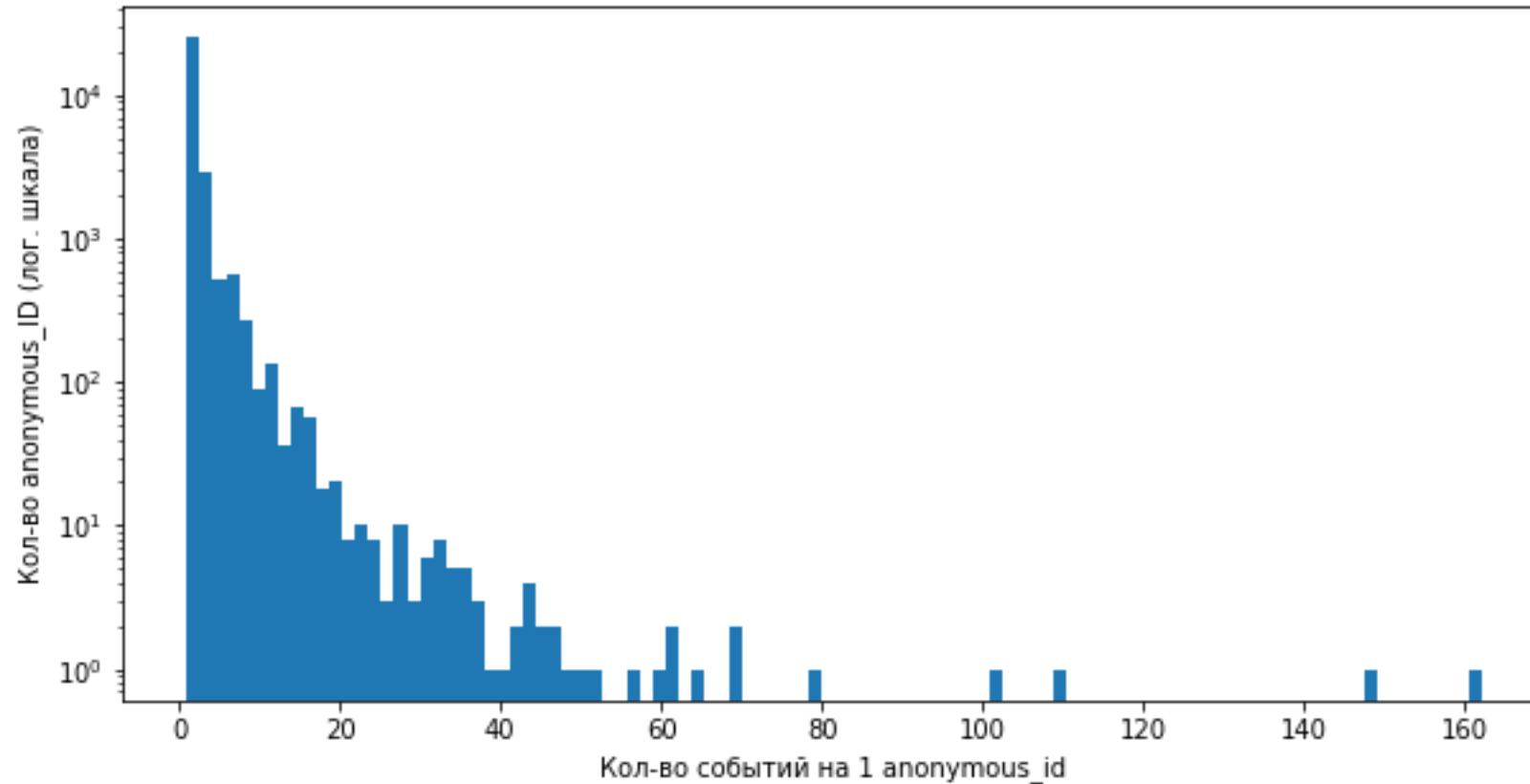
Если это одно из событий текущего эксперимента, которое не описано в задаче, то скрипт разработаем так, чтобы легко добавить его.

Вывод3. Отсутствие *session_id* вносит подозрения на корректность эксперимента. Необходимо знать номер сессии и понимать в каком эксперименте она участвует.

Скрипт python3 для расчетов: <https://clck.ru/UqZw7>

Shop_Selected

Распределение кол-ва событий Shop_Selected на 1 anonymous_id



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

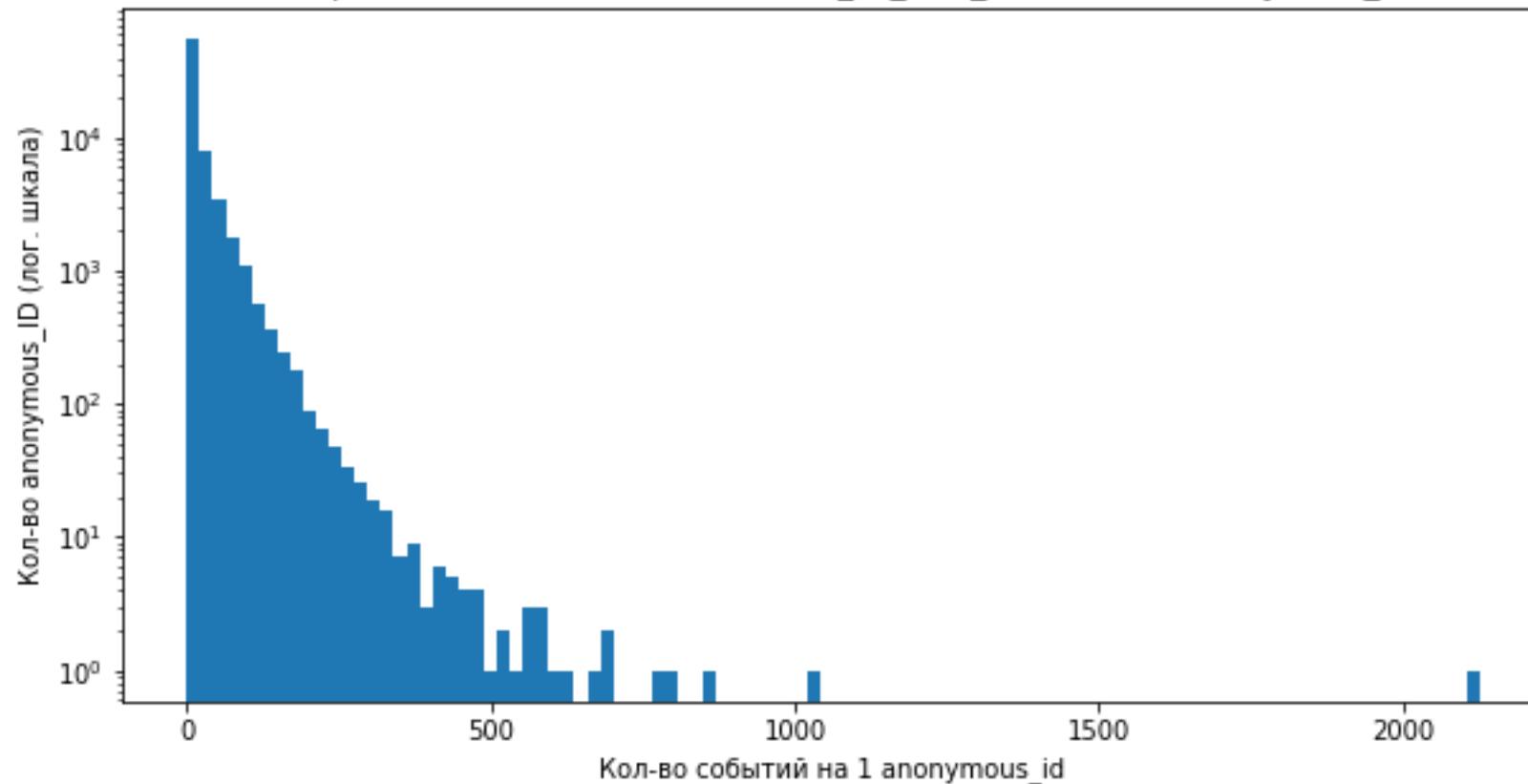
99% перцентиль = 32

Будем считать кол-во событий на один anonymous_id >32 аномальными

Удалим аномальные anonymous_id

Add_To_Cart_Clicked

Распределение кол-ва событий Add_To_Cart_Clicked на 1 anonymous_id



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

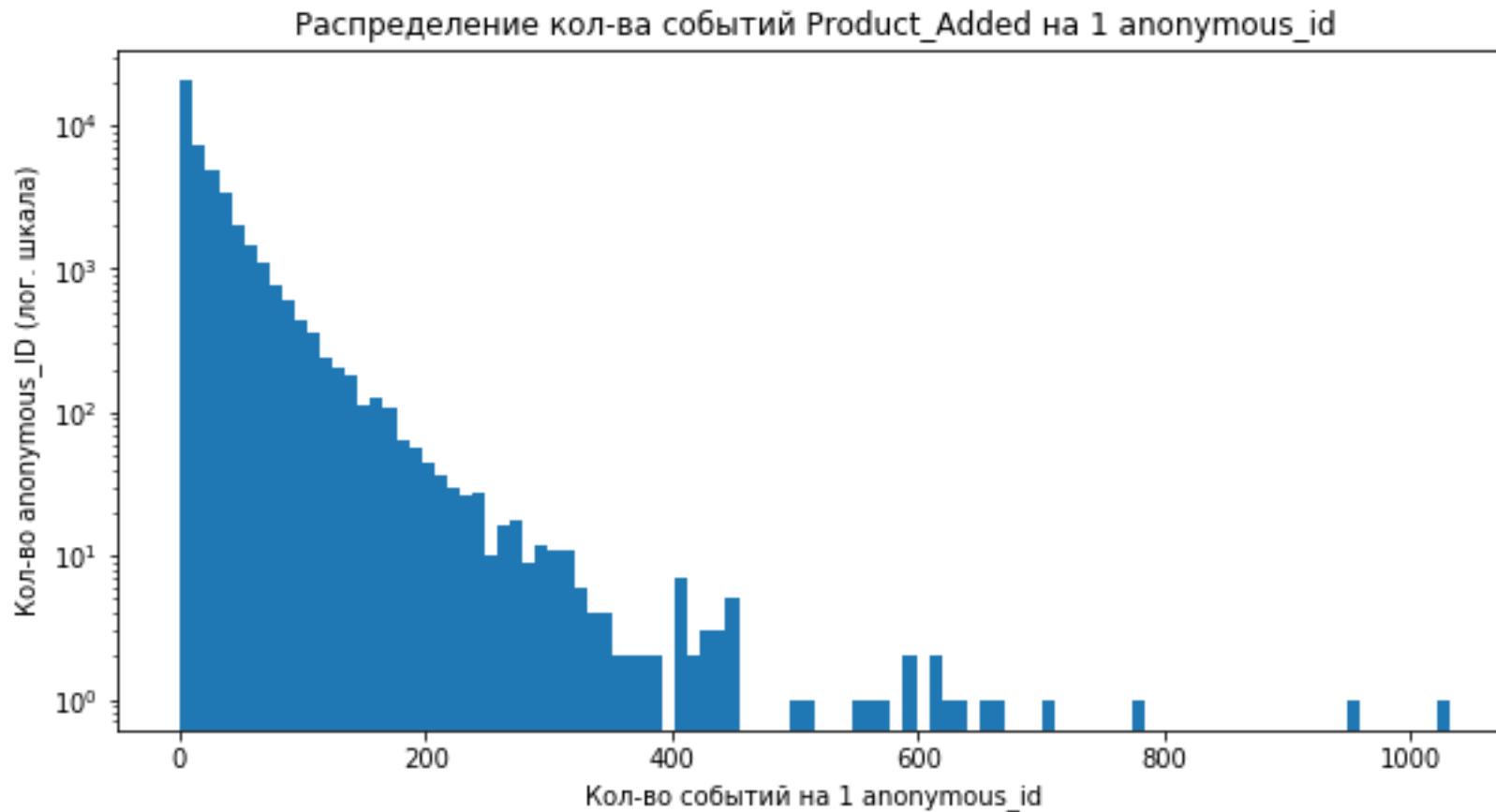
Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 155

Будем считать кол-во событий на один anonymous_id > 155 аномальными

Удалим аномальные anonymous_id

Product_Added



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

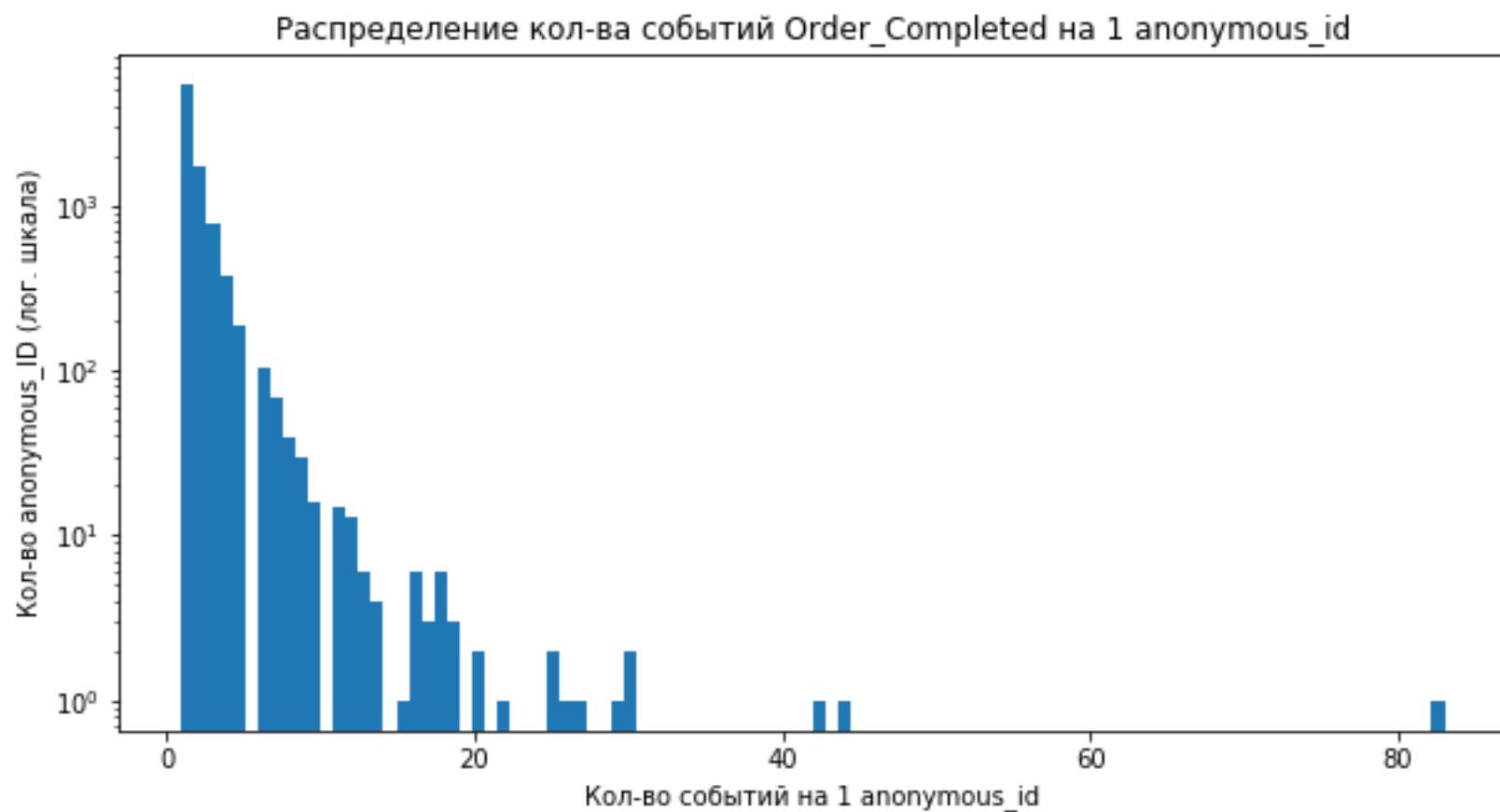
Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 175

Будем считать кол-во событий на один anonymous_id > 175 аномальными

Удалим аномальные anonymous_id

Order_completed



Алгоритм построения

1. Расчет кол-ва событий на 1 Anonymous_ID
2. Построение гистограммы (для демонстрации)
3. Ось Y логарифмическая

Вывод: есть аномальные выбросы кол-ва событий на один anonymous_id

99% перцентиль = 9

Будем считать кол-во событий на один anonymous_id >9 аномальными

Удалим аномальные anonymous_id

Гипотеза по аномально большим событиям

- Большое количество действий, часть из них завершены заказами, выполняются преимущественно на десктопных устройствах или планшетах с высокой регулярностью. (см. табл. на след. слайде)
- Можно предположить, что эти заказы выполняют профессиональные закупщики – рестораны, общепиты, мелкие ретейлеры.

Гипотеза по аномально большим событиям.

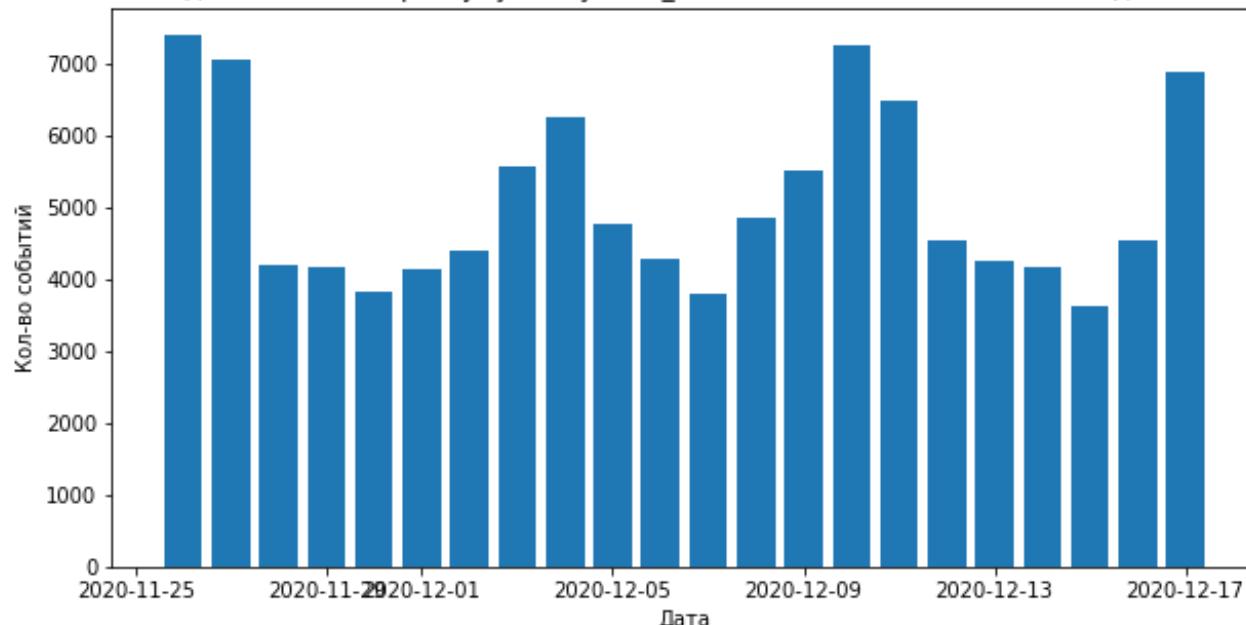
Аргументы

Распределение по устройствам пользователе с аномальным кол-во событий Product_Added

Тип устройств	Кол-во устройств во всей выборке	Уникальны пользователи, которые добавляют товар	Доля от общего числа
desktop	184130	373	0.002026
mobile	298372	58	0.000194
tablet	6549	10	0.001527

1. Доля ID с аномально большими кол-вами заказах в десктопах значительно больше, чем в мобильниках

Кол-во событий "добавления в корзину" у anonymous_id с аномально большим кол-вом добавлений в корзину



2. кол-во действий от этих пользователей растет в нестандартное для физлиц дни недели
- 26.11 - чтв
 - 27.11 - птн
 - 03.12 - чтв
 - 04.12 - птн
 - 10.12 - чтв
 - 11.12 - птн
 - 17.12 - чтв

Дополнительно

- Исключены anonymous_id, которые попали в обе группы

```
select anonymous_id, count(distinct group)  
from AB_Test_Hit  
group by anonymous_id  
having count(distinct group) =1
```

- Исключены anonymous_id, которые были в контроле, но источник в таблицах «Address Change Initiated» или «Shop_Selected» у них “landing”
- Исключены anonymous_id, которые были в teste, но источник в таблицах «Address Change Initiated» или «Shop_Selected» у них “add_product” или “address_modal”

Эксперимент

Гипотезы для проверки

- **H1.** Измененный LP + путь выбора адреса конвертирует в Main_Page также, как первоначальный LP
- **H2.** Конверсия в тестовой группе с LP в завершенный заказ больше контрольной
- **H3.** Конверсия в тестовой группе с LP в добавление товара больше контрольной
- **H4.** Тестовый LP дает процент отказов такой же, как контрольный LP
- **H5.** В тестовой группе конверсия с Main_Page в добавление товара больше, чем в контрольной
- **H6.** В тестовой группе конверсия с Main_Page в завершение заказа больше, чем в контрольной

Все гипотезы исследовать в «сессиях»

Методология

1. Исключить все аномалии и сомнительные моменты (по source)
2. Построить сессии
 - Берем только тех, кто по нашим данным впервые на сайте
 - Сессией считаем все события которые находятся по времени между двумя посещениями LP
 - Время завершения последней сессии в рамках anonymous_id = последнему timestamp в эксперименте.
3. Определить сессии с целевыми действиями
4. Рассчитать конверсии
5. Применить t-test для процентных величин (конверсий)

Частые обозначения

- **cnt_total_control** – общее кол-во событий в контроле
- **cnt_target_control** – целевое кол-во событий в контроле
- **CR_control** – конверсия в контроле
- **cnt_total_test** - общее кол-во событий в тесте
- **cnt_target_test** - общее кол-во событий в тесте
- **CR_test** - конверсия в контроле
- **Bounce_r_c/Bounce_r_t** – процент отказов в контроле/тесте
- **Bounce_sess_c/Bounce_sess_t** – сессии с отказами контр./тест

Результаты 1

Конверсия с LP в Main_Page:

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
217706	70599	0.324286	17431	809	0.046412	True

H1. гипотеза 1 стат.значимо провалилась. В контрольной группе конверсия с LP в главную страницу в несколько раз больше тестовой

Конверсия с LP в заказ:

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
217706	1172	0.005383	17431	80	0.00459	False

H2. Результат статистически не различим. Гипотеза 2 не подтвердилась

Результаты 2

Конверсия с LP в добавление товара

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
217706	7504	0.034469	17431	365	0.02094	True

Н3. гипотеза 3 статистически значимо опровергнулась. В тестовой группе конверсия с LP в добавление товара стат. значимо меньше контрольной

Конверсия из LP в добавление товара зависит от типа девайса

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest	cnt_total_control
desktop	82991	4812	0.057982	7619	335	0.043969	True
mobile	131472	2605	0.019814	9515	25	0.002627	True
tablet	3242	86	0.026527	296	5	0.016892	False

Хуже всего ситуация изменилась в мобильных устройствах

Результаты 3

Отказы LP

Cnt_total_control	Bounce_sess_c	Bounce_r_c	cnt_total_test	Bounce_sess_t	Bounce_r_t	ttest
217706	146974	0.675103	17431	16312	0.935804	True

Н4. гипотеза 4 статистически значимо опровергнулась. В тестовой группе отказы стат. значимо больше контрольной

Device_type	Cnt_total_control	Bounce_sess_c	Bounce_r_c	cnt_total_test	Bounce_sess_t	Bounce_r_t	ttest
desktop	82991	58533	0.705293	7619	6938	0.910618	True
mobile	131472	86182	0.655516	9515	9089	0.955229	True
tablet	3242	2265	0.698643	296	284	0.959459	True

Хуже всего ситуация изменилась в мобильных устройствах

Результаты 4

Конверсия с Main_Page в добавление товара:

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
70599	7504	0.10629	809	365	0.451174	True

H5. гипотеза 5 статистически значимо подтвердилась. В тестовой группе конверсия с Main_Page в добавление товара в несколько раз больше контрольной

Конверсия с Main_Page в завершение заказа:

cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
70599	1172	0.016601	809	80	0.098888	True

H6. гипотеза 6 статистически значимо подтвердилаась. В тестовой группе конверсия с Main_Page в заказ в несколько раз больше тестовой

Результаты 5

Стат. значимое различие

Конверсия из Main_Page в добавление товара по ретейлерам*

Retailer_id	cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
29	27804	3092	0.111207	296	107	0.361486	True
1	27670	4327	0.156379	1007	346	0.343595	True
57	16745	1900	0.113467	263	88	0.334601	True
93	3531	116	0.032852	8	2	0.250000	False
33	2169	165	0.076072	4	1	0.250000	False

The diagram shows two red arrows originating from the 'CR_control' values in the table. One arrow points from the value '0.111207' (row 29) to a blue bracket on the right labeled '+25%'. Another arrow points from the value '0.113467' (row 57) to a blue bracket labeled '+19%'.

Прирост CR в тесте над контролем разный от retailer_id

Идея1 (для следующего эксперимента): добавление ретейлеров, проработка Main_Page ретейлера и развитие доставки влияет на общую конверсию

Идея2 (для следующего эксперимента): изменение последовательности по-разному оказало влияние на разных ретейлеров (надо больше данных для проверки)

*Построены сессии, где разделителем начала и конца Main_Page.

Результаты 6

Конверсия из Main_Page в добавление товара по устройствам*

Device_type	cnt_total_control	cnt_target_control	CR_control	cnt_total_test	cnt_target_test	CR_test	ttest
mobile	66018	3397	0.051456	328	33	0.100610	True
desktop	38555	6823	0.176968	1310	521	0.397710	True
tablet	1446	106	0.073306	13	6	0.461538	False

Прирост конверсии в desktop больше мобильных устройств

*Построены сессии, где разделителем начала и конца Main_Page.

Выводы1

В тестовой группе:

- Конверсия из LP в заказ стат. значимо не изменилась
- Конверсия из LP в Main_Page стат. значимо ухудшилась
- Конверсия из LP в добавление товара стат. значимо ухудшилась
- Процент отказов LP стат. значимо вырос
- Конверсия с Main_Page в добавление товара выросла
- Конверсия с Main_Page в заказ выросла

Выводы2

Измененная последовательность позволила улучшить часть пути пользователя с Main_Page в добавление товаров и заказ,

но глобально результат был поглощен независимым экспериментом с Лендингом

В итоге конверсии в заказ теста и контроля не различимы

Следующие шаги

1. Проговорить с Оунером, что все события, которые были исключены, не относятся к эксперименту (это надо было обсудить до расчетов ☺)
2. Объяснить, что необходимо провести повторно эксперимент без изменения LP.
Предварительно устранить все выявленные ошибки
 1. рассчитать размеры выборок
 2. установить трекер сессий по всем страницам
 3. создать справочник {сессия: эксперименты}
 4. устранить технические проблемы с перекидыванием пользователей в разные группы
3. Если эксперимент повторно подтвердится (или повторно провести уже нельзя), то раскатать в прод новую последовательность с тестовым LP
4. Протестировать разные LP сократить отказы, особенно на мобильном
5. Протестировать разные адресные модалки. Увеличить CR в Main_Page
6. Подключать новых партнеров и развивать доставку, чтобы исключить тех, пользователей, к которым не может быть доставлен товар. Тем самым, увеличивать конверсию из адресной модалки в Main_Page