

Предсказание прибыльности пользователя на основе истории покупок и возвратов

Решетняк А.О.
Дружинин А.О.
Терещенкова Л.Ю.
Грушко А.А.



Рассмотренные гипотезы

- Влияние скорости покупки на прибыль
- Влияние браузера на суммы возвратов
- Влияние количества платформ на процент возвратов

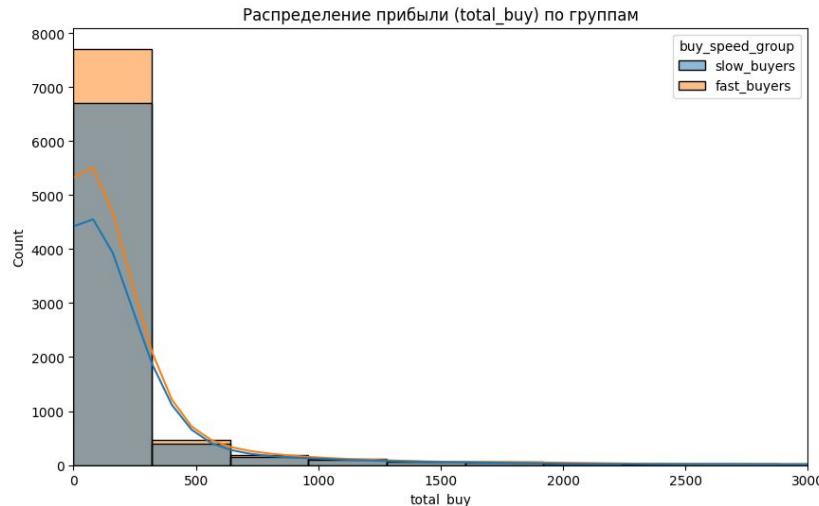
[Ссылка на дашборд](#)



Гипотеза 1: влияние скорости покупки на прибыль.

В ходе анализа было выявлено влияние скорости покупок на величину прибыли.

Вся выборка была поделена на две части по медианному значению скорости покупки (разность даты регистрации первой покупки) - slow_buyers и fast_buyers.



Влияние скорости покупки на прибыль

По критерию Манна-Уитни **p-value ≈ 0.006**. Значит, разница в прибыли от двух групп по скорости прибыли статистически значима. Пользователи, покупающие быстро, приносят больше прибыли.

Дополнительно было выяснено, что скорость принятия решения о первой покупке (быстро или медленно) не влияет на объем возвратов (**p-value ≈ 0.8**).

Рекомендации для бизнеса:

- Предлагайте "быстрые" триггеры для первой покупки: скидки на первый заказ, бесплатная доставка, ограниченные по времени предложения.
- Минимизируйте шаги между регистрацией и покупкой. Внедрите "быструю покупку" для зарегистрированных пользователей.
- Сразу после регистрации предлагайте персонализированные рекомендации (если есть таргет) или популярные товары, чтобы сократить время на принятие решения.

Гипотеза 2: влияние браузера на суммы возвратов

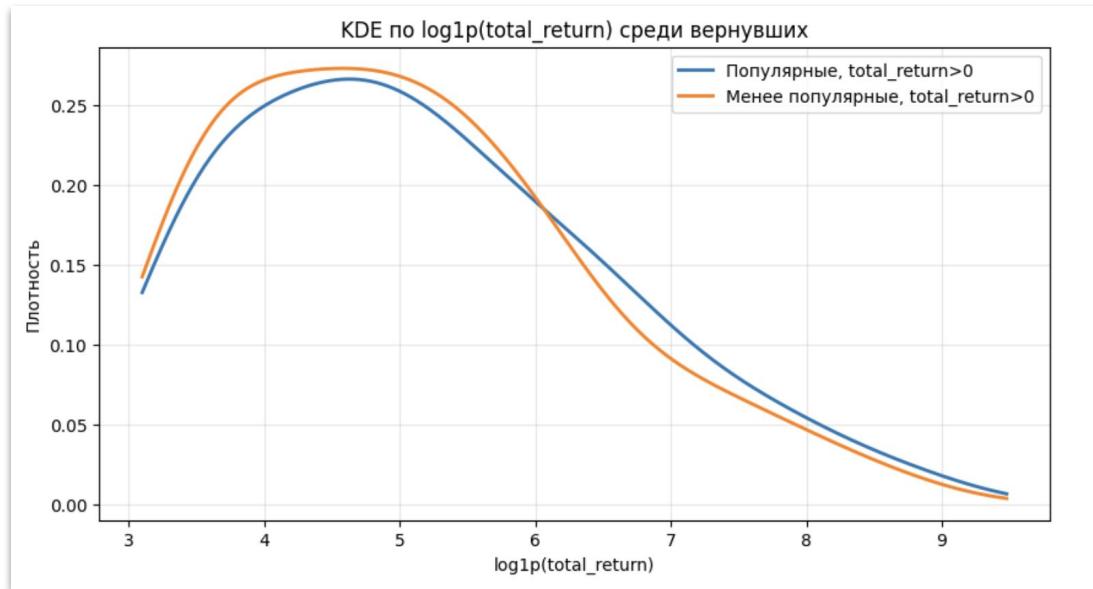
В ходе анализа было выявлено влияние покупок в популярных браузерах на возвраты.

Было выделено 2 самых популярных браузера: **browser_02** и **browser_03**.

Группа с наиболее популярными браузерами сравнивалась с группой с менее популярными браузерами.

Влияние браузера на суммы возвратов

Для визуализации распределений сумм возвратов использован KDE-plot по $\log1p(\text{total_return})$ среди пользователей, совершивших возврат.



Визуализация показывает небольшие различия в форме распределений, в частности более тяжёлый правый хвост у пользователей популярных браузеров.

Влияние браузера на суммы возвратов

Поскольку сумма возвратов — количественный показатель, групп по браузерам 2, данные в них независимы и распределены ненормально, для проверки гипотезы был использован U-критерий Манна-Уитни.

Тест показал **p-value ≈ 0.04**, что ниже уровня значимости 0.05, а это значит, что среди пользователей, совершивших возврата, распределение величины возвратов статистически значимо отличается между популярными и менее популярными браузерами.

Рекомендация: подробнее проанализировать это влияние, и учитывать влияние популярности браузеров при активации пользователя. В зависимости от популярности браузера, пользователи могут возвращать товары на более крупные суммы.

Гипотеза 3: влияние количества платформ на процент возвратов

Наблюдение: При анализе клиентов, чей процент total_return больше 30% от total_buy была замечена аномальная активность.

Предположение: Убыточные пользователи используют значительно больше уникальных платформ (устройств/браузеров) для входа, чем обычные клиенты. Это может указывать на работу ботов или скриптов.

Подтверждение аномалии:

Статистический тест подтвердил ($p<0.05$), что пользователи, приносящие компании убыток, используют аномально высокое количество платформ для входа.

Медиана у обычных клиентов: 2 платформы.

Медиана у убыточных: 4 и более, со средним значением около 10.

Причина убытков:

Живой человек редко использует более 3-4 уникальных устройств. Высокие показатели platform_num (вплоть до 50+) в сочетании с массовыми возвратами указывают на работу ботов или мошеннических скриптов, которые перебирают устройства для обхода лимитов или накрутки.



Практические рекомендации:

Настроить систему антифрода на автоматическую пометку пользователей, у которых `platform_num > 5` за короткий период.

Блокировка: Пользователем с `platform_num > 10` и отрицательным балансом покупок (возвраты > покупок) усложнить процедуру возврата или требовать дополнительную верификацию (например, подтверждение по SMS) перед оформлением заказа/возврата.

Чистка базы: Провести аудит топ-50 пользователей по количеству платформ и заблокировать явных ботов, чтобы не искажать маркетинговую аналитику в будущем.

**Спасибо
за внимание!**