

Результаты A/B теста sbermarket.ru

Перенос шага выбора адреса на лендинг

Конверсия тестовой группы выросла, bounce rate - упал

В бой стоит пускать вариант показанный тестовой группе - с перенесенным шагом выбора адреса.

	Тестовая группа	Контрольная группа	Изменение
Конверсия	15,1%	12,7%	2,39%
Отказы	46,3%	54,6%	-8,2%

Все результаты статистически значимы: доверительные интервалы оценок конверсий не содержат нуля

Комментарии по полученным результатам

Конверсия рассчитана по событиям добавления в корзину после включения в группу, отказы - как отсутствие хотя бы 1 события с момента включения, в группы.

При этом при более строгом подсчете - конверсии с учетом всей воронки, а отказов, только по следующему шагу после landing view результат противоположный - конверсия падает на 4 пп, отказы растут на 7,2 пп.

Однако, скорее всего эти результаты обусловлены ошибками сбора данных событий.

Дизайн эксперимента. Предложения по улучшению

- | | |
|----------------------------|---|
| Гипотеза | <ul style="list-style-type: none">Формулировка гипотезы не четкая (понятие “избежать кривых сценариев”). Если все равно выбрали метрики лучше сразу так и сказать “Выбор адреса до этапа выбора магазина повысит конверсию в добавление в корзину без повышения отказов с лендинга” |
| Изменения | <ul style="list-style-type: none">Внесены дополнительные изменения в лендинг (появился блок “введите адрес” и вместо экрана “выберите магазин” выводится блок “Партнеры”, что само по себе может влиять на полученные результаты). Можно было оставить все как есть, а при клике на магазин выводить поп-ап выбора адреса (если для указанного адреса доступен выбранный магазин сразу переход в каталог) - это не добавляет лишних шагов, позволяет избежать ситуации, описанной в качестве проблемы и сам лендинг остается максимально без изменений. |
| Пользователи | <ul style="list-style-type: none">Для контрольной группы не хватает события на успешный выбор адреса.Ничего не сказано про критерии отбора пользователей. Лучше включать только новых пользователей, которые ранее не были на сайте или на лендинге.В дизайне нет оценки необходимого количества пользователей и принципа (пропорции) деления на группы. |
| Ожидания и действия | <ul style="list-style-type: none">Нет предположений об ожидаемых изменениях конверсии (вырастет/упадет, на сколько) и четкого плана действий (достаточно понятно, но все же стоит детализировать, что будем делать если конверсия и bounce останутся на примерно тех же уровнях) |

Анализ данных. Проблемы

- Событий AB Test Hit значительно больше, чем событий landing viewed. Есть большая доля пользователей с AB Test Hit без Landing Viewed (282833 хитов, 56% от общего числа). Возможно, событие Landing Viewed отправляется только после полной загрузки страницы, а AB Test Hit - до.
- Есть пользователи, попавшие сразу в 2 группы (56310 хитов, 11% от общего числа). Проблема механизма распределения.
- В эксперимент включены вернувшиеся пользователи, которые уже имели опыт взаимодействия с сайтом (судя по датам событий Main Page viewed), при этом такие пользователи включались в обе группы. Ошибка дизайна эксперимента.
- Есть существенный дисбаланс при делении пользователей на группы - 11% (в тестовой) к 89%. Это не привело к проблеме статистической значимости, но этого можно было избежать рассчитав нужный объем выборок.
- Непонятные параметры событий. Для тестовой группы значения address, add_product, add_alcohol, not_in_delivery_zone, которых по дизайну быть не должно.

Анализ данных. Рекомендации/типичные ошибки

- Проконтролировать распределение трафика между группами и рассчитать необходимый размер групп, для получения статистически значимых результатов. Особенно если трафик делится не 50/50 по группам.
- Проконтролировать работу механизма распределения, случайность и равномерность распределения пользователей (в данном случае проблем нет)
- Убедиться что в ходе эксперимента нет существенного отличия трафика от обычного (не проводятся, например, какие-то особые маркетинговые активности)
- В ходе анализа необходимо очистить данные от дублей, противоречивых и неполных данных.
- Проверять результаты на статистическую значимость

Ссылки

Расчеты выполнялись в excel

Metrics.xlsx - метрики и оценка стат. значимости.

Groups stats.xlsx - проверка равномерности по ос, устройствам и тд, а также распределения по группам.