# **ФИО:** Иванов Владислав Михайлович

**Описание:** У вас есть 2 дня для прохождения теста. Ответы принимаются в любом удобном формате (Word, .pdf и т.п.). Целью является правильное (а в задачах 4 и 5 еще и подробное) решение задач.

**Рекомендация:** Вначале прочтите внимательно все задачи и начните решать наиболее для вас простые.

**Задача 1.** Конверсия из клика в регистрацию составляет 4%, конверсия из регистрации в лид (целевое действие пользователя) 30%. Какова вероятность, что из 60 кликнувших пользователей мы получим хотя бы 1 лид?

Рассчитаем вероятность получения лида

p = конверсия\_рег \* конверсия\_дейст = 0.04 \* 0.3 = 0.012

Рассчитаем вероятность не успеха получения лида при одном действии

q = 1 – P = 1 – 0.012 = 0.988 –

Тогда вероятность не получить ни одного лида при 60 кликов равна:

q^60 = 0.988^60 = 0.485

Вероятность получить хотя бы 1 лид при 60 кликов будет равна:

P = 1- 0.485 = 0.515 или 51.5%

**Задача 2.**

Вы подбрасываете кубик (6 граней), после чего у вас есть две возможности:

1. Взять себе сумму $, равную выпавшему на кубике числу (выпало 3 – получаете 3$)

2. Отклонить результат первого броска и подбросить кубик второй раз. После чего уже взять себе столько $, сколько выпало во второй раз (второй бросок отклонять уже нельзя)

**Какую стратегию нужно выбрать, чтобы максимизировать ожидаемую прибыль (при каких значениях первого броска нужно перебрасывать?)**

**Решение:**

Ожидаемая прибыль первого броска равна: P1 = 1/6 \* (1+2+3+4+5+6) = 3.5

Если при первом броске выпадает число меньше 3, мы должны перебросить кубик второй раз.

Если выпадает 4 5 или 6, то отказываемся перебрасывать.

**Чему равна ожидаемая прибыль при этой стратегии (математическое ожидание)?**

**Решение:**

Ожидаемая прибыль согласно стратегии:

P2 = 3/6 \* 3.5 + 1/6 \* (4+5+6) = 4.25 (> 3.5)

**Задача 3.** Есть таблица table с транзакциями пользователей.

В ней есть поля id, user\_id, datetime, amount. Нужно написать SQL-запрос, который считает сумму первых транзакций по пользователю.

SELECT

user\_id,

SUM(amount) AS total\_amount

FROM (

SELECT

user\_id,

amount,

ROW\_NUMBER() OVER (PARTITION BY user\_id ORDER BY datetime) AS row\_num

FROM table

) AS subquery

WHERE row\_num = 1

GROUP BY user\_id

**Задача 4.**

Перед вами стоит задача оценки окупаемости трафика который приходит по СРА (оплата за целевое действие, в нашем случае первый депозит). У вас есть база со всеми нужными вам данными за несколько лет. Вы закупили 100 игроков заплатили за каждого из них по $10. Вы наблюдаете за ними 7 дней. Вам нужно построить прогнозную модель, которая предскажет через сколько времени вернутся вложенные деньги и посчитать LTV игрока.

**Для построения прогнозной модели проведем следующие расчёты показателей:**

**Рассчитаем** жизненную ценность клиента LTV за 7 день по нашей группе игроков

LTV = AVG(Gross Profit) x Lifetime

Lifetime = 1/(1- Retention)

**Рассчитаем** коэффициент рентабельности инвестиций

ROI = 100 \* (LTV \* Количество приобретенных клиентов - Затраты на привлечение клиентов) / Затраты на привлечение клиентов

Далее на основе полученного значения можем оценить время возвращения вложенных средств.

Как пример, если показатель будет равняться 100%,то период окупаемости наступит через неделю, если 50% - то через две недели.

Как вы будете использовать имеющиеся у вас исторические данные?

LTV будем считать на основе валовой прибыли, поэтому потребуются данные о переменных расходах, которые растут пропорционально продажам. (комиссии, временные саппорты и т.п.)

Какие еще параметры вам необходимо знать для построения модели?

1. Идентификатор игрока, который позволит отслеживать его активность и возврат вложенных денег.

2. Дата внесения депозита, позволит строить когорты для точного расчета LTV

3. Сумма внесения депозита

4. Активность игрока, для расчета Retention

5. Рыручка

Какие фичи (переменные) трафика вы будете использовать для прогнозирования?

Количество уникальных посетителей сайта за каждый день наблюдения (7 дней).

Количество пользователей, внесших депозит за каждый день наблюдения.

Количество игроков, сделавших первую покупку в течение первых 7 дней после регистрации.

Суммарный доход от этих новых игроков. Доход по дням.

Количество обращений в службу поддержки за каждый день наблюдения.

Время, проведенное каждым игроком на сайте за каждый день наблюдения.

Среднее время между первой покупкой и второй покупкой каждого игрока.

Количество игр, сыгранных каждым игроком за каждый день наблюдения.

Какие 3 фичи вам кажутся наиболее важными?

*Количество игроков, сделавших первую покупку в течение первых 7 дней* после регистрации. Эта переменная является ключевой для прогнозирования времени возврата вложенных денег, так как покупка является действительным проявлением заинтересованности в продукте и готовности заплатить за его использование.

*Средняя выручка на одного игрока в течение первых 7 дней после регистрации*. Для оценки эффективности маркетинговой кампании и прогнозирования времени возврата вложенных денег важно знать, сколько денег тратит каждый игрок на сайте.

*Доля игроков, которые вернулись на сайт после первой покупки.* Эта переменная показывает, насколько успешно сайт удерживает своих пользователей после первоначальной покупки

Какие метрики вы выведете в модель в качестве расчетных?

Средняя валовая прибыль в течение первых 7 дней после регистрации.

Доля игроков, которые совершают повторные покупки в течение первых 7 дней после первой покупки.

Затраты на привлечение клиентов

Количество приобретенных клиентов

**Задача 5.**

Представьте, что вы аналитик в крупном интернет-магазине. За последний месяц факт продаж оказался на 15% ниже прогноза и вам необходимо выяснить, в чем причина. Опишите, как бы вы решали эту задачу: какие гипотезы вы бы стали проверять, каким образом и какие данные вам для этого потребуются.

Весь процесс поиска причин снижения продаж в интернет-магазине разделим на шаги:

1. **Проведем сбор данных** за последний месяц, так же за аналогичный месяц предыдущего года. Идеально собрать данные за предыдущие три года, чтобы рассмотреть общую динамику продаж магазина.

*Данные могут содержать:* время совершения покупки, id продукт/товар, цена, количество продуктов в корзине, категория продукта, скидки и акции, каналы продаж, купленный трафик, маркетинговые компании, регион продаж, клиентская информация (возраст, пол, история, предпочтения), возвраты и отмены, сбои работы сайта.

1. **Сформируем гипотезы:** На основе анализа данных, сформулируем несколько гипотез, которые могут объяснить причину снижения продаж.

3.1 Гипотеза: Снижение продаж связано с изменением цен на товары.

Действия: Сравнить средние цены на товары в последний месяц с предыдущим периодом. Проанализировать, какие товары и категории продуктов имеют наибольшее изменение в цене.

3.2 Гипотеза: Снижение продаж связано с недостаточной рекламной активностью.

Действия: Проанализировать, какие каналы рекламы приводят к наибольшему количеству продаж, а какие работают некорректно, либо не эффективно. Сравнить с последним месяцем.

3.3 Гипотеза: Снижение продаж связано с сезонными факторами.

Действия: Сравнить продажи в текущем месяце с продажами в предыдущих летних месяцах. Сравнить аналогичный период предыдущих годов.

3.4 Гипотеза: Снижение продаж связано с конкурентами.

Действия: Изучить конкурентные акции и маркетинговые стратегии, которые могли повлиять на продажи.

3.5. Гипотеза: Снижение продаж связано с недовольными клиентами.

Действия: Изучить отзывы и жалобы клиентов за последний месяц. Оценить уровень удовлетворенности клиентов и выяснить, есть ли какие-либо общие тенденции или проблемы, которые могут влиять на продажи.

3.6 Гипотеза: Снижение продаж связано с техническими сбоями.

Действия: Изучить метрики стабильности работы сайта, регистрации покупателей, проверить возможность совершения оплаты и т.п.

1. **Проведем анализ данных**, для выявления явных тенденций и аномалий.
   1. Анализ продаж по времени: оценим продажи по дням, неделям и месяцам. Определим пики и интервалы времени с минимальным и максимальным количеством продаж. Будем использовать графики и диаграммы для визуализации данных.
   2. Анализ продаж по категориям товаров: оценим объем продаж и доходность каждой категории товаров за последний месяц, проведем сравнение с предыдущими месяцами. Категории с наибольшим долей в общих продажах требуется оценить какие-либо негативные изменения.
   3. Анализ клиентского поведения: изучим данные о покупателях, такие как история покупок и показатели удовлетворенности. Можно использовать сегментацию клиентов и провести анализ для каждого клиента.
   4. Анализ эффективности маркетинговых кампаний: оценим данные по электронным рассылкам (email, mobilepush, sms, webpush), рекламным акциям, скидочным предложениям. Оценим эффективность, определим какие каналы работают, на какие покупатели лучше всего реагируют. Рассмотрим изменения за последний месяц.
   5. Анализ отзывов и оценок клиентов: рассмотрим отзывы клиентов, проставленные оценки, рассмотрим динамику изменения. Можем прочитать комментарии.
   6. Анализ работоспособности платформы сайта магазина: Изучим данные о производительности и стабильности платформы сайта магазина. Проанализируем метрики, такие как время загрузки страниц, время отклика сервера, количество ошибок или проблемных запросов, приводящие к потере продаж.

Анализ данных поможет выявить тенденции и аномалии и предоставляет информацию для дальнейшего принятия решений и определения причин снижения продаж.

1. **ДОПОЛНИТЕЛЬНО:**

* Соберем обратную связь других отделов
* Разработаем план действий
* Мониторинг и оценка.