

Создание плана повышения retention платформы IT Resume

Удержание клиентов – важная часть в построении успешного бизнеса. Для его измерения используется метрика Retention Rate, которая и будет рассмотрена в данной работе.

Платформа IT Resume представляет сервис по решению задач с несколькими видами подписок: Free (бесплатная с ограниченным доступом) и Premium VIP (платная на 1 месяц/ без ограничений по времени).

Необходимо провести анализ активности клиентов, чтобы разработать план повышения Retention. Для этого посчитаем метрики и выдвинем следующие гипотезы:

1. Для определения когорт и выбора оптимального промежутка отслеживания Retention просмотрим периоды регистрации и заходов пользователей, найдем среди них минимальное и максимальное значение;

2. Сравним даты регистрации пользователей с первым заходом на платформу. Выдвинем гипотезу, что эти даты совпадают;

3. Произведем расчет Rolling Retention с учетом дат в п.1. Т.к. IT Resume платформа с подписочной системой, то данная метрика хорошо отразит долю пользователей, которые заходили в день n и после него;

4. Рассмотрим N-day Retention для первых 7 дней, чтобы понять, как в общем падает активность от момента первого захода на платформу. Предполагаем, что активность в первые дни максимальна, а к 7 дням начинает падать. При этом пользователи продолжают совершать заходы еще длительное время, т.к. у них есть подписка (и это покажут расчеты в п.3).

5. План повышения retention.

При разработке плана повышения retention будут даны рекомендации с выдвижением гипотез, которые не могут быть подтверждены или опровергнуты в связи с ограниченным доступом к базе данных.

1. Определение когорт и промежутка отслеживания

Произведем объединение пользователей по когортам, которые указываются на основании месяца регистрации. Также посмотрим, сколько человек в одной когорте и есть ли необходимость по данному месяцу производить дальнейшие расчеты (рисунок 1).

Запрос SQL:

```
select
    count(distinct id) as cnt,
    date_trunc('month', date_joined)::date as date
from
    users
where true
    and id >= 94
    and {{date_joined}}
group by
    date_trunc('month', date_joined)
```

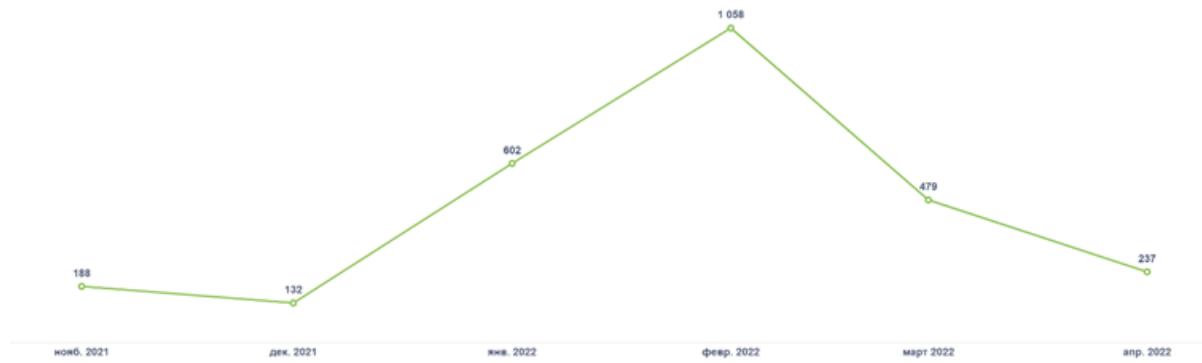


Рисунок 1 – Распределение зарегистрированных пользователей по месяцам

На основании рисунка 1 делаем выводы, что данных достаточно для выделения когорт по всем имеющимся месяцам.

Далее рассмотрим периоды активных заходов на платформу (рисунок 2).

Запрос SQL:

```
select
    min(entry_at::date) as min,
    max(entry_at::date) as max
from
    userentry u
where
    user_id >= 94
```

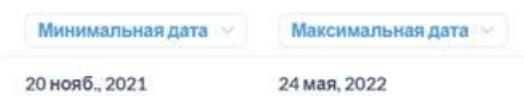


Рисунок 2 – Период активных заходов

Данные по заходам пользователей имеются с ноября 2021 года по май 2022. Расчет Rolling retention произведем по следующим дням: 0, 1, 3, 7, 14, 30, 60, 90. Стоит отметить, что для «поздних» когорт retention будет равен 0 не из-за отсутствия активности, а из-за отсутствия данных и невозможности подсчета.

2. Анализ зарегистрированных пользователей

Рассмотрим дату первого захода пользователей на платформу и сравним ее с датой регистрации. Ранее была выдвинута гипотеза, что эти данные совпадают (рисунок 3).

Запрос SQL:

```
with all_date as (
    select
        b.user_id,
        b.date as joined_date,
        a.date as entry_date,
        a.date - b.date as diff
    from (
        select
            user_id,
```

```

min(entry_at)::date as date
from
    userentry
group by
    user_id
order by
    user_id,
    date
) as a
right join (
    select
        id as user_id,
        date_joined::date as date
    from
        users
    where true
        and { {date_joined} }
) as b
on a.user_id = b.user_id
where
    b.user_id >= 94
),
type_diff as (
    select
        case when diff = 0 then 'День регистрации'
            when diff > 0 then 'Последующие дни'
            else 'Не заходил'
        end as entry,
        user_id
    from
        all_date
)
select
    entry,
    count(user_id) as cnt
from
    type_diff
group by
    entry

```

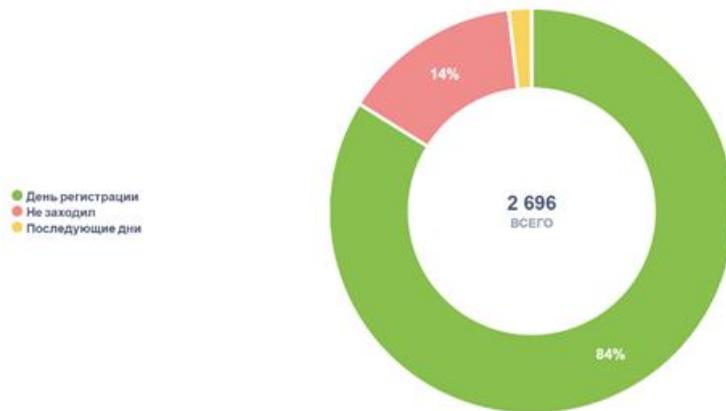


Рисунок 3 – Распределение пользователей по первому заходу на платформу после регистрации

Из рисунка 3 видно, что гипотеза частично подтвердилась. Основная часть пользователей заходит на платформу в день регистрации, небольшая часть – спустя время, а 14% пользователей ни разу не заходили на платформу.

3. Расчет Rolling Retention

Произведем расчет Rolling Retention с учетом ранее полученных данных (рисунок 4).

Запрос SQL:

```
with all_active as (
    select
        users.id,
        date_joined,
        entry_at,
        extract(day from (entry_at - date_joined)) as day_e,
        to_char(date_joined , 'YYYY-MM') as cohort
    from
        users
    left join userentry
    on users.id = userentry.user_id
    where true
        and users.id >= 94
)
```

```

        and {{date_joined}}}
)
select
    cohort,
    round(count(distinct id) filter (where day_e >= 0) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "0 (%)",
    round(count(distinct id) filter (where day_e >= 1) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "1 (%)",
    round(count(distinct id) filter (where day_e >= 3) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "3 (%)",
    round(count(distinct id) filter (where day_e >= 7) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "7 (%)",
    round(count(distinct id) filter (where day_e >= 14) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "14 (%)",
    round(count(distinct id) filter (where day_e >= 30) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "30 (%)",
    round(count(distinct id) filter (where day_e >= 60) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "60 (%)",
    round(count(distinct id) filter (where day_e >= 90) * 100.0 / count(distinct id) filter
    (where day_e >= 0), 2) as "90 (%)"
from
    all_active
group by
    cohort
order by
    cohort

```

Когорта		0 (%)	1 (%)	3 (%)	7 (%)	14 (%)	30 (%)	60 (%)	90 (%)
2021-11	100	71,19	65,54	62,71	60,45	53,11	41,24	29,94	
2021-12	100	50,41	43,8	42,15	35,54	31,4	23,14	13,22	
2022-01	100	41,77	35,54	29,72	25,3	15,26	8,84	4,62	
2022-02	100	29,18	20,99	16,27	12,68	9,43	5,61	1,23	
2022-03	100	30,6	24,82	18,31	13,98	8,19	1,2	0	
2022-04	100	43,75	35,58	28,37	21,63	10,1	0	0	

Рисунок 4 – Rolling Retention

На рисунке видно, что есть когорты, сильно отличающиеся низкой активностью уже в первые дни. Наиболее «удачными» оказались когорты 2021 года. Несмотря на то, что количество новых пользователей в этих месяцах была не самой высокой (рисунок 1), наблюдается стабильная активность. Значения с «0» не оцениваем, т.к. данных нет.

4. N-day Retention

Проведем также анализ активностей пользователей в течение 7 дней от первого захода на платформу (Рисунок 5 – N-day Retention).

Запрос SQL:

```
with a as (
    select
        user_id,
        to_char(min(entry_at) over (partition by user_id), 'YYYY-MM') as cohort,
        entry_at,
        min(entry_at) over (partition by user_id) as first_active,
        extract('day' from entry_at - min(entry_at) over (partition by user_id)) as n_day
    from userentry
    where true
        and user_id >= 94
        and { {entry_at} }
),
b as (
    select
        cohort,
        n_day,
        round(count(distinct user_id) * 100.0 / max(count(distinct user_id)) over (partition by cohort), 2) as retention
    from
        a
    where
        n_day <= 7
    group by
        cohort,
        n_day)
select
```

cohort as “Когорта”,
 min(case when n_day = 0 then retention end) as “0(%)”,
 min(case when n_day = 1 then retention end) as “1(%)”,
 min(case when n_day = 2 then retention end) as “2(%)”,
 min(case when n_day = 3 then retention end) as “3(%)”,
 min(case when n_day = 4 then retention end) as “4(%)”,
 min(case when n_day = 5 then retention end) as “5(%)”,
 min(case when n_day = 6 then retention end) as “6(%)”,
 min(case when n_day = 7 then retention end) as “7(%)”
 from
 b
 group by
 cohort

Когорта		0(%)	1(%)	2(%)	3(%)	4(%)	5(%)	6(%)	7(%)
2021-11	100	17,14	16	13,14	8	11,43	11,43	8,57	
2021-12	100	9,92	9,92	9,92	9,92	4,96	8,26	4,96	
2022-01	100	14,66	7,94	8,76	6,31	6,72	4,68	5,7	
2022-02	100	9,07	6,38	4,7	3,7	3,58	2,58	2,24	
2022-03	100	9,11	7,19	7,67	4,32	5,04	4,08	2,64	
2022-04	100	12,79	12,21	13,66	8,14	6,1	6,4	5,52	
2022-05	100	14,13	10,95	6,36	6,36	5,3	5,3	4,24	

Рисунок 5 – N-day Retention

По рисунку 5 видно, что активность после первого захода очень сильно падает во всех когортах. Но как показал Rolling Retention (рисунок 4), пользователи продолжают пользоваться платформой в течение определенного времени. Стоит отметить, что в тех когортах, где Rolling Retention «проседает», уже в первую неделю наблюдается низкая активность среди пользователей. Поэтому задача данной работы, разработать план по повышению retention, является важной и актуальной.

5. План повышения retention

Ранее был проведен анализ активных пользователей на основании зарегистрированных (рисунок 3). Есть 14% пользователей, которые не пользовались платформой после регистрации. Было бы интересно узнать о них больше.

Для начала нужно определить, являются ли они действительно потенциальными клиентами или же их регистрация на платформе была «случайностью». Для этого предлагается отправить на электронную почту письмо с просьбой пройти опрос, при этом всем прошедшим мы подарим промокод на скидку n%. Заголовок письма может быть следующим: «С Вас – опрос, с нас – скидка 20%», «Ответь на 2 вопроса и получи скидку 20%».

Далее описать ситуацию: «Мы заметили, что Вы давно к нам не заходили. Поделитесь с нами Вашим мнением и получите скидку на подписку n%». Скидку можно сделать разного размера, выставить ограничение по времени действия скидки и т.д.

Сам опрос высыпается спустя определенное время после регистрации (это может быть 10, 15, 30 дней и т.д.) и состоит из вопросов:

1. Планирует ли пользователь дальше заходить на платформу (да/нет);
2. Что не устроило на платформе (цена, наполненность сайта, отсутствие мотивации к решению и т.д.).

Во-первых, с помощью опроса сразу можно определить, кто является потенциальным клиентом. Если пользователь вообще не проявил активность и не прошел опрос, то его относим к категории «случайных». С такой категорией нет смысла работать. Если люди проходят опрос, то они попадают в категорию потенциальных клиентов и задачей здесь является их привлечение на платформу.

Во-вторых, на основании ответов на второй вопрос делаются выводы о том, где можно провести изменения для повышения активности пользователей.

Далее рассмотрим пользователей, которые уже активно пользуются платформой. Выдвигаем гипотезу, что каждый пользователь использует 1-2 языка программирования для решения задач. Данную гипотезу нельзя подтвердить или опровергнуть, т.к. работа проводится по ограниченной базе. Допустим, это действительно так. И большинство использует SQL + Python/Java/C и т.д. Или же использует только один язык. Стоит расширить платные подписки и ввести возможность выбрать 1 или 2 языка за меньшую стоимость по тарифам.

Также проведем анализ подписок, какую покупают чаще всего: на месяц или без ограничений по времени. Если окажется, что большинство пользователей отдает предпочтение подписке на месяц, то тариф без ограничений по времени, но с меньшей суммой может сподвигнуть их на покупку. При этом в полном тарифе Premium VIP должны оставаться явные преимущества по сравнению с тарифами по ограниченному количеству языков (о них будет далее).

Рассмотрим цель пользователей платформы. Выдвинем гипотезу, что имеется две основные категории (не сможем подтвердить или опровергнуть, т.к. нужно проводить опрос):

1. Пользователи, желающие просто прокачать навыки;
2. Пользователи, желающие найти работу благодаря пройденному материалу.

Вторая категория должна иметь четкое понимание, что задачи/тесты помогут пройти собеседования и устроиться на работу. На данный момент нет подробного описания таких случаев, кроме «рекрутеры мониторят сайт».

Нужно давать больше четкой информации, например: «Пользователь на собеседовании решал задачу *ссылка на задачу*. Благодаря хорошей подготовке он получил предложение о работе в компанию К. Оформите подписку со скидкой n% и оцените свои скиллы». Либо же как рекрутер нашел сотрудника через IT Resume. Здесь также подойдет актуальная рассылка/ пост в блоге, который будет мотиватором для решения задач.

Рассмотрим первую категорию пользователей, которая хочет прокачать навыки и не имеет цели трудоустройства. Здесь вводится рейтинг среди всех пользователей (может быть как общий, так и с фильтром по языкам программирования). Чем больше сложных задач решено, тем выше место. Важный момент: в Premium VIP (самый дорогой тариф) возможность находиться в рейтинге входит в стоимость, а для всех остальных участие является платным. Этот рейтинг имеют возможность видеть работодатели, т.е. те, кто будет выше, автоматически быстрее привлечет их внимание. И хотя не все пользователи имеют цель трудоустройства, возможно, их заинтересует конкретное предложение. При этом должна быть возможность «скрыть себя» из рейтинга, видимого для работодателей (тариф Premium VIP).

Пользователям, которые решают задачи для прокачки навыков, а также испытывают спортивный интерес – быть лучше других, необходимо делать рассылку, если появляется новая задача или другой пользователь обходит их в рейтинге. Здесь также подойдет введение акции: если пользователь приглашает друзей или коллег посоревноваться в решении задач, то оба получают скидку.

Подводя итог вышесказанному, хочется отметить следующее: чтобы вводить изменения, необходимо сначала изучить мотивацию пользователей. После этого разрабатывать и внедрять в работу новшества, описанные в данной работе.