# Что такое RFM-анализ

В работе аналитика часто встречается задача - провести RFM-анализ. Но иногда это задание вызывает трудности, поскольку студентам не так часто объясняют что такое RFM-анализ и с чем его едят, хотя это крайне полезный метод. Давайте разберемся!

# Что это такое?

RFM-анализ - это оценка клиентов по их транзакциями и активности. RFM-анализ позволяет сегментировать клиентов по частоте, сумме и давности, и выявлять «китов», за которых нам еще предстоит побороться, и тех, на кого можно не тратить время и силы. Получается, мы можем определить ценность клиента, а это ключевой момент для любого бизнеса.

#### RFM - расшифровка:

- Recency (давность): как давно клиент совершал заказы.
- Frequency (частота покупок): сколько раз клиент совершал заказы.
- Monetary (деньги): на какую сумму клиент совершал заказы.

RFM-анализ активно используется в маркетинге, чтобы лучше понять клиентов и определить, какие из них нуждаются в большей внимательности, каких можно привлечь новыми продуктами или услугами.

# Как это работает?

Суть RFM-анализа в том, что мы разделяем всех клиентов на группы, опираясь на то, как давно они сделали последнюю покупку, как часто покупали и насколько большой была сумма их заказов.

По каждому из этих признаков мы выделяем по три равные группы. Затем присваиваем каждой группе числовое обозначение от 1 до 3.

По давности заказа (recency):

- 1 свежие клиенты;
- 2 сравнительно недавние клиенты;
- 3 давние клиенты.

#### По частоте покупок (frequency):

- 1 постоянный покупатель;
- 2 покупает нечасто;
- 3 покупает очень редко (единичные заказы).

## По сумме покупок (monetary):

- 1 большая сумма;
- 2 средняя сумма;
- 3 маленькая сумма.

## Примеры:

- Например, клиент «333» покупал давно, один раз и на маленькую сумму. И скорее всего, нам не стоит за него бороться.
- Но при этом клиентов «313», «312», «311» обязательно нужно попытаться вернуть - это наши уходящие постоянные клиенты с разным чеком.
   Отправляем им реактивационные письма, запрашиваем обратную связь, предлагаем бонусы, программы лояльности.
- Или пользователь «111»: покупает часто, на большую сумму и последняя покупка была недавно. Это наши лучшие клиенты. Таким клиентам следует уделять особое внимание, например, предлагать им эксклюзивные продукты или услуги, бонусы и скидки.

# Что с этим делать?

С каждой группой можно строить отдельные коммуникации: давать им разную рекламу и делать разные email-рассылки. Например, группе постоянных VIP-клиентов дать платиновую карту, а пользователям, которые давно не покупали — мотивирующую скидку, и настроить на них таргетированную рекламу.

Границы давности, частоты и суммы покупок вы определяете самостоятельно. То есть, для какого-то бизнеса 10 000 руб. - это много для 1 клиента, а для какого-то - почти ничего. Или частота покупок на сайте авиабилетов будет существенно отличаться от частоты покупок в магазине «у дома». Поэтому здесь можете проявить фантазию, отталкиваясь от вашего бизнеса.

Например, мы предлагаем разбивать значения всех признаков (давность, частота, сумма) по перцентилям и в качестве границ выбирать 0%-33%, 33%-66%, 66%-100%. Это стандартная разбивка - мы выработали ее для себя на практике, а дальше ее можно уже корректировать на конкретных данных.

Главное - помнить основной принцип: мы всегда должны «тянуть» каждого клиента в группу повыше. То есть лучших клиентов мы стараемся сделать еще лояльней и продать им еще больше за счет удержания, средних «довести» до лучших, а уходящих и почти потерянных — вернуть. Потому что вернуть клиента всегда легче, чем привлечь нового!

Ниже - пример реального RFM-анализа, который сделал один из студентов нашего Симулятора «Аналитик данных» в качестве итоговой работы в конце модуля по SQL. Здесь и код, и графики, и бизнес-выводы.

# RFM-анализ клиентской базы аптечной сети

#### Цели и задачи

Основной мотив исследования: скоро сезон активных продаж спадет и выручки в аптеке начнут падать. У компании есть механизм СМС-рассылок, но хотят провести персонализированную рассылку, чтобы не потратить бюджет впустую.

Для этого было проведено исследование клиентской базы аптечной сети с помощью RFM-анализа, который позволил классифицировать покупателей на основании истории их покупок и были сформированы рекомендации, какой группе покупателей что предлагать.

#### Об RFM-анализе

RFM-анализ — метод анализа, позволяющий сегментировать клиентов по частоте и сумме покупок и выявлять тех, которые приносят больше денег.

Аббревиатура RFM расшифровывается:

Recency — давность (как давно пользователи что-то покупали)

Frequency — частота (как часто они покупают)

Monetary — деньги (общая сумма покупок).

По каждому из этих признаков мы выделяем по три равные группы. Затем присваиваем каждой группе числовое обозначение от 1 до 3

Давность последней покупки (Recency)	Частота покупок (Frequency)	Сумма покупки (Monetary)
1 - недавние клиенты	1 - частые	1 - большая
2 - редкие нерегулярные	2 - редкие	2 - средняя
3 - давние, возможно потерянные	3 - очень редкие	3 - маленькая

После разделения людей на группы получаем максимум 27 сегментов.

Анализ помогает определить клиентов, которые покупают часто и на большие суммы, клиентов, которые регулярно возвращаются за небольшими покупками, и клиентов, которые давно ничего не покупали

Таргетирование рассылок на определенные группы клиентов дает гораздо большую конверсию рекламы, нежели одинаковая не персонализированная реклама.

### Структура базы данных

Для исследования использовалась таблица **Bonuscheques** (БД Аптека), которая содержит информацию о транзакциях по бонусным картам **card** (т.е.это постоянные клиенты, они оставили свои телефоны)

Если в момент покупки касса была в оффлайн-режиме, то вместо номера карты записывается зашифрованная последовательность символов. В таком случае номер карты силами этой базы данных никак восстановить нельзя. Поэтому в исследование взяты покупки, которые можно идентифицировать.

Также из исследования исключены покупатели, сделавшие одну покупку – им нет смысла предлагать персональную скидку для повышения среднего чека.

datetime - дата и время совершения транзакции

summ - сумма чека

#### Выбор метрик для анализа

Основная сложность RFM-анализа — определить границы сегментов, потому что универсальных рекомендаций по этому поводу нет.

В данном исследовании для того, чтобы оценить распределение покупателей, были рассчитаны средние, медианы и перцентили, а также построен boxplot.

Среднее – самый популярный статистический показатель, но чувствительный к выбросам, которые смещают среднее значение и могут ввести в заблуждение. Медиана показывает центральное значение в выборке.

В нашем случае средняя сумма покупок 5018 руб., медианная сумма 2708 руб. Значения сильно отличаются, и мы можем предположить наличие выбросов (отдельные клиенты покупали на очень большие суммы).

Среднее количество покупок 5,7, медианное количество - 3. Можем предположить, что у аптечной сети много разовых покупателей, которых хорошо бы сделать постоянными лояльным клиентами.

Перцентили позволяют нарезать слоями массив данных и понять, на какие группы делятся элементы. Таким образом, покажут изменения с большей точностью, но спрячут выбросы.

Для большей точности было сделано распределение по децилям.

Для того, чтобы убедиться в правильности расчетов и посмотреть визуально, было сделано распределение по квартилям (построен боксплот)

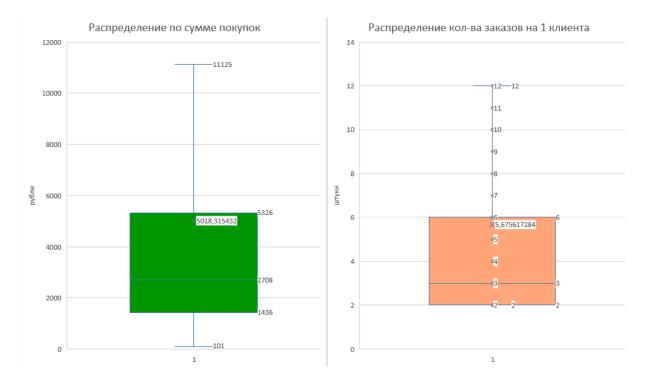
И подсчитаны кумулятивные сумма и количество заказов, т.к.закон Парето про 80/20

В исследовании по каждому из трех параметров разделили покупателей на 3 группы. В результате получились вот такие сегменты:

Маркер	Давность последней покупки (Recency)	Частота покупок (Frequency)	Сумма покупки (Monetary)
1	Покупка 26 и менее дней назад	7 и более покупок	6346 р.и больше
2	84 и менее	3-6 покупок	2136-6345 руб.
3	Покупка более 84 дней назад	2 покупки (очень редко)	2135 р. и меньше

Почему выбраны именно эти границы сегментов?

Всего период, который взят для анализа - полтора года (можно выбрать другой временной период с помощью фильтра на дашборде). Этого отрезка времени достаточно, а) чтобы можно было сделать ретроспективный анализ, б) не тратить сила на совсем уж давних клиентов, потому что клиенты могли исчезнуть потому что переехали, сменили работу, и им стало неудобно, или сменили телефон, и бюджет будет тратиться впустую на рассылку по неактуальным номерам.



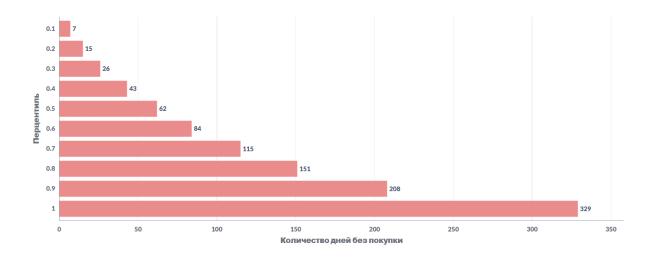
#### Давность последней покупки (Recency)

Из распределения по децилям видно, что 30% клиентов - те, у кого последняя покупка была 26 и меньше дней назад. Т.е.это постоянные клиенты и новые, которые были недавно и аптечная сеть хочет, чтобы они пришли снова. Это самая "свежая" группа 1.

Медиана (62 дня) и среднее (85 дней). Т.е.30% клиентов относим в среднюю категорию.

Те, кто не приходил в аптеку более 85 дней (почти 3 месяца) дней попадают в группу 3, потому что это либо в принципе редкие посетители аптек (например, молодые люди), либо те, кому стало территориально неудобно, либо ушедшие к конкурентам.

Распределение количества дней после последней покупки ("пропущенных" дней)



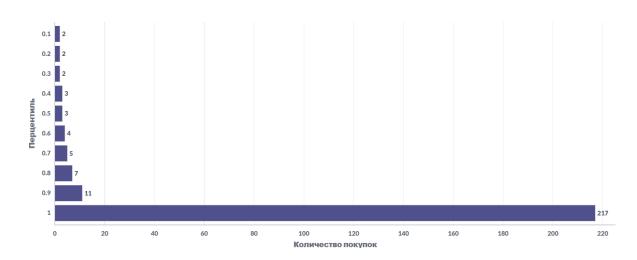
# Частота покупок (Frequency)

Достаточно много разовых клиентов (33%), принесших 11,5% от общего количества посещений. Они самые "холодные", поэтому группа 3.

Есть клиенты, которые за полтора года приходили 7 и более раз (23% от общего количества покупателей, как видно из распределения по децилям. Их количество покупок 51,65% от общего количества покупок (накопительный итог)). Такие лояльные и постоянные клиенты получают категорию 1.

Медианное количество покупок - 3. Это основной пул клиентов (44 %). Среднее 6. Накопительный итог - 89%, т.е.это подтверждает, что большая часть покупателей сделала от 3 до 6 покупок за период. Категория 2.

#### Распределение количества покупок на одного клиента



#### Сумма покупки (Monetary)

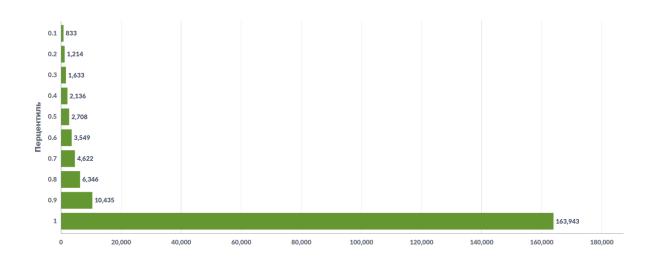
Из соотношения медины и среднего можно понять наличие выбросов, поэтому к среднему надо относиться с осторожностью. Из распределения и накопительного итога мы видим, что 60% продаж приносят 20% покупателей, которые покупали на сумму 6346 и более руб. Категория 1.

Категория 2 - основная масса клиентов, которая покупает на сумму от 2136-6345 руб. Вместе с категорией 1 они приносят 90% продаж (накопительный итог), составляя 60% клиентов (четвертый дециль)

В эти границы попадает медиана (2708 р.) и среднее (5018 р.)

Категория 3 - оставшиеся 40% клиентов. которые приносят всего 10% продаж.

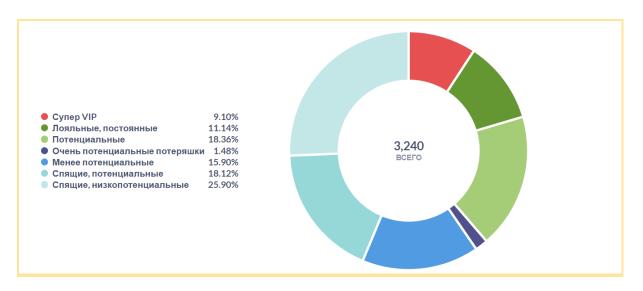
#### Распределение суммы покупок на одного клиента



Некоторые сегменты есть смысл объединить, если они очень маленькие по количеству покупателей или сходны по стратегии работы с ними.

У отдела маркетинга тоже очень много задач и написать 27 вариантов работы с группами, в которых 10-50 человек - и затратно, и может возникнуть путаница при отправке сообщений, ну и вряд ли эти 27 вариантов будут как-то концептуально отличаться между собой, т.е.может получиться "работа ради работы", а не результата.

Поэтому покупатели были объединены в 7 логических групп:



Групп	Кол-во	Описание	Группа логическая
· Py	клиенто 🕆	▼	r pyrina norm reekan
111	295	недавно, часто, чек высокий	Супер VIP

Групп	Кол-во клиенто ▼	Описание	Группа логическая
112	120	недавно, часто, чек средний	Лояльные, постоянные
121	39	недавно, 3-6 покупок, чек высокий	Лояльные, постоянные
122	202	недавно, 3-6 покупок, чек средний	Лояльные, постоянные
	361		

Групп	Кол-во клиенто ▼	Описание	Группа логическая
311	48	давно, часто, чек высокий	Очень потенциальные потеряшки

Группэ	Кол-во	Описание	Группа логическая
ı pyılı	клиенто 🕆	OTINCATIVE T	труппа логическая
211	146	нерегулярно, часто, чек высокий	Потенциальные
212	84	нерегулярно, часто, чек средний	Потенциальные
221	50	нерегулярно, 3-6 покупок, чек высокий	Потенциальные
222	244	нерегулярно, 3-6 покупок, чек средний	Потенциальные
231	5	нерегулярно, редко, чек высокий	Потенциальные
232	66	нерегулярно, редко, чек средний	Потенциальные
	595		

Групп	Кол-во клиенто ▼	Описание	Группа логическая
113	5	недавно, часто, чек низкий	Менее потенциальные
123	114	недавно,3-6 покупок, чек низкий	Менее потенциальные
131	3	недавно, редко, чек высокий	Менее потенциальные
132	56	недавно, редко, чек средний	Менее потенциальные
133	148	недавно, редко, чек низкий	Менее потенциальные
213	8	нерегулярно, часто, чек низкий	Менее потенциальные
223	181	нерегулярно, 3-6 покупок, чек низкий	Менее потенциальные
	515		

Группа	Кол-во	Описание	Группа логическая
Групп	клиенто 🔻	-	труппа логическая
312	50	давно, часто, чек средний	Спящие, потенциальные
321	56	давно, 3-6 покупок, чек высокий	Спящие, потенциальные
322	317	давно, 3-6 покупок, чек средний	Спящие, потенциальные
331	7	давно, редко, чек высокий	Спящие, потенциальные
332	157	давно, редко, чек средний	Спящие, потенциальные
	587		

Групп	Кол-во клиенто ▼	Описание	Группа логическая
233	186	нерегулярно, редко, чек низкий	Спящие, низкопотенциальные
313	3	давно, часто, чек низкий	Спящие, низкопотенциальные
323	223	давно, 3-6 покупок, чек низкий	Спящие, низкопотенциальные
333	427	давно, редко, чек низкий	Спящие, низкопотенциальные
	839		

# Дашборд

 $\frac{https://metabase.simulative.ru/public/dashboard/2b75517c-7582-4d31-b0f2-4a4c021c7d43}{\#refresh=3600}$ 

RFM-анализ. Таблица итоговая с присвоением категории каждому клиенту



## SQL-запрос

Ниже SQL-запрос, написанный для СУБД PostgreSQL и содержащий специфические конструкции фильтров для BI-системы Metabase, в которой был построен дашборд (ссылка и скриншот - в предыдущем блоке).

```
with identify as (
                                     --выбираем клиента, даты покупок,
СУММЫ
                select card, datetime, summ
                       from bonuscheques
                       where card like '200%'
                       and {{datetime}}
                            ),
sum_cnt as (
                                                 --группируем по
клиенту, убираем тех, у кого 1 покупка
            select card,
                sum(summ) as sum_by_card,
                count(summ) as cnt_orders_by_card
            from identify
            group by card
            having count(summ)>=2
recency as (
            select card,
            last_value(datetime) over (partition by card rows between
unbounded preceding and unbounded following) as max_order_date --дата
последней покупки
                 from identify
```

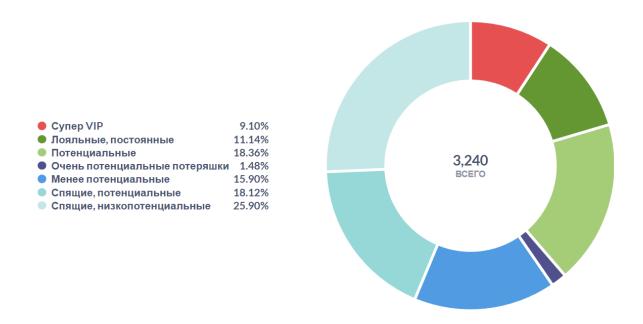
```
),
for_rfm as (select recency.card as "Клиент",
                    max(max_order_date::date) as "Дата покупки",
                    make_date(2022,6,10)-max(max_order_date::date) as
"Дней без покупок",
                          --беру 10.06 2022 как допущение, что у меня
свежая выгрузка (последние данные от 09.06.22)
                    max(cnt_orders_by_card) as "Количество покупок",
--просто обернула в функцию, т.к.есть groupby
                    max(sum_by_card) as "Сумма покупок"
            from recency
            right join sum_cnt on sum_cnt.card=recency.card
--соединяю так, чтобы из таблицы с датой последней покупки ушли
1-покупочные клиенты
            group by recency.card
            ),
rfm as (
        select "Клиент",
                "Дата покупки",
                "Дней без покупок",
                (case when "Дней без покупок"<=26 then '1'
                    when "Дней без покупок" <=84 then '2'
                    else '3' end) as "R",
                "Количество покупок",
                (case when "Количество покупок">=7 then '1'
определяю категорию для каждого из компонентов анализа
                    when "Количество покупок" >=3 then '2'
                    else '3' end) as "F",
                "Сумма покупок",
                (case when "Сумма покупок">=6346 then '1'
                    when "Сумма покупок">=2136 then '2'
                    else '3' end) as "M"
        from for_rfm
        ),
gr_name as (
            select "Клиент",
                             "Дата покупки",
                             "Дней без покупок",
                             "R",
                             "Количество покупок",
                             "Сумма покупок",
```

```
"M",
                            concat("R", "F", "M") as "RFM" -- вывожу
группу покупателей для RFM анализа
            from rfm
            order by "RFM"
select "Клиент",
        "Дата покупки",
        "Дней без покупок",
        "R",
        "Количество покупок",
        "Сумма покупок",
        "M",
        "RFM".
        (case when "RFM" in ('112', '121', '122') then 'Лояльные,
постоянные'
            when "RFM" in ('211', '212', '221', '222', '231', '232')
then 'Потенциальные'
            when "RFM" in ('113', '123', '131', '132', '133', '213',
'223') then 'Менее потенциальные' -- формирую группы и называю их
            when "RFM" in ('312', '321', '322', '331', '332') then
'Спящие, потенциальные'
            when "RFM" in ('111') then 'Cyπep VIP'
            when "RFM" in ('311') then 'Очень потенциальные
потеряшки'
            else 'Спящие, низкопотенциальные' end) as "Название
группы"
from gr_name
```

#### Выводы

В результате проведенного RFM-анализа клиентской базы аптечной сети были сформированы 7 групп покупателей для рассылки смс-сообщений с целью увеличить чек и частоту покупок.

Распределение покупателей получилось удобным и похожим на распределение по правилу Парето, которое уважаемо и применимо в работе в условиях ограниченных ресурсов, человеческих и денежных. А ресурсы ограничены всегда)



Группы достаточно большие, чтобы не распылять силы, но при этом объединены по своему поведению.

#### Фокусы глобально:

- 1. Увеличить частоту посещения аптеки
- 2. Увеличить средний чек

#### Это можно сделать, проведя клиентов по маркетинговой воронке лояльности:

- "теплых" и "потеплевших" "горячими"
- 🚼 "горячих" склонить к покупке и сделать VIPами

## Теперь подробнее:

1) создать широкий охват людей в верхней части воронки: разбудить две группы спящих и привлечь новых клиентов.

Как мы увидели из RFM-анализа - редких покупателей у аптечной сети многовато, и фокус на перевод случайных покупателей в регулярные - самый важный, потому что, даже при низком чеке, покупатели - это выручка и будущие лояльные клиенты.

#### Спящие потенциальные (18,1%)

Они покупали на достойные суммы, но куда-то пропали. Это одна из самых важных групп, потому что их можно конвертировать в группу активных покупателей.

Надо про нас напомнить и дать какую-то заманушку:

предложения в стиле "до даты X двойные бонусы"; начислены дополнительные бонусы по карте и они сгорят числа Y; акция по снижению стоимости на препараты ABC до числа X".

Т.е.надо прийти в аптеку до определенного времени

Спящие низкопотенциальные (25,9%)

Давние, по сути разовые, покупатели, покупали по-мелочи.

С ними могло случиться все что угодно, например, территориальные изменения (ну не пойдет человек в аптеку далеко от дома или работы). Мог смениться номер - тогда бюджет, потратится в никуда. Могла сильно не понравиться сама аптека. Или это молодые люди богатырского здоровья, которые крема-шампуни-презервативы покупают на Вайлдбериз.

Риски: бюджета надо много, а результата может быть мало. Если кто-то и вернется - то окупит ли это затраченные на них средства?

Возможности: все же получить постоянных покупателей. В этой группе много 2-х разовых клиентов, по сути - случайных. Но, если пришел 2 раза, а не один - значит не такой уж он и случайный? Может, он живет рядом, но ходит в другую аптеку и его лояльность можно развить?

Может клиент не знает, что аптека - это не только лекарства, но и парафармация? Аптеке больше доверия, чем непонятному продавцу с маркетплейса.

Клиент считал аптеку дорогой - вот смс об акции, что цена доступная. Клиент переехал и перестал ходить в аптеку - так есть другие аптеки сети рядом.

Можно предлагать акции на парафармацию, скидку при регистрации в приложении, приветственные баллы при оформлении карты.

Для холодных клиентов и, т.к. лек.препараты и парафармация - продукт специфический, к его рекламе нужно подойти более фундаментально, чтобы объяснить потенциальным клиентам его ценность. Например, рассылка новостей или научно-популярных статей. Такая реклама создает меньше ощущения продажности, чем контент для маркетинга, поэтому она может быть более привлекательным способом оставаться на связи с клиентами и стимулировать посещение сайта, чтение обзоров продуктов или обращение за дополнительной информацией.

Могу предположить, что компания подключает на этом этапе и другие инструменты, кроме смс-рассылки (заказы через личный кабинет на сайте, поисковую оптимизацию для интернет-запросов, ретаргетинг и поп-апы с предложениями скидок, соцсети и другие способы повышения осведомленности).

Отдельно хотела бы отметить (это не относится непосредственно к поставленной задаче, но аналитик может помочь) важность удержать вновь пришедших клиентов:

удобство сайта и личного кабинета, простота регистрации и оформления покупок и т.д. Узнать об этом можно можно из опросов для постоянных покупателей.

#### 2) Три группы потенциальных "тёплых" клиентов

#### Потенциальные (18,3%)

Платежеспособные клиенты, которые покупают нерегулярно. Основная цель - чтобы приходили чаще.

Нужно периодически напоминать о товарах компании, рассказывать о новых продуктах, отправлять уведомления об акциях. Хорошо - когда надо прийти в аптеку до определенного времени (бонусы сгорят, акция до числа X и тд.).

#### Менее потенциальные (15,9%)

Покупают на меньшие деньги, чем Потенциальные. Цель - увеличить сумму и частоту их покупок. Подойдут акции "купи 2, получи третий в подарок", о снижении цены, начисление бонусных баллов на какую-либо категорию товаров.

#### Очень потенциальные потеряшки (1,5%, 48 человек)

Покупали давно последний раз, но часто и на большие суммы.

Основная задача — реактивировать, снова сделать их супер лояльным, как-то о них позаботиться (персонализированные супер-акции). Эта группа маленькая, но она заслуживает отдельного внимания, т.к.может перейти в категорию Супер VIP и внести хороший вклад в прибыль. И супер-акция обойдется недорого, т.к.клиентов в этой группе мало.

#### 3) Лояльные клиенты

#### Супер VIP (9,1%)

Это абсолютные лидеры среди клиентов. Они приносят существенную долю выручки компании + могут быть неплохим источником сарафана. Это категория «главные фанаты».

Хорошо было бы прислать им что-то супер персонализированное, заказать под них что-то, что они регулярно покупают, позаботиться о них, дать бы им VIP карту и так далее. Сделать что-то, чтобы у них случилась реакция «нифига себе» и они побежали рассказывать об аптечной сети друзьям. Рекомендации повышают вероятность того, что люди попробуют приобрести, и, как правило, приводят к более быстрой конверсии.

#### Лояльные, постоянные (11,1%)

Это тоже группа лидеров. Уже приносят аптечной сети основной доход и могут прирасти еще.

Им можно предлагать новинки, приглашать подписаться на соцсети, проводить с ними опросы.

Цель работы с группой - повысить их средний чек за счет того, что они узнают о новых для них товарах.

Категориям лояльных клиентов не надо очень часто делать рассылку, чтобы она не воспринималась как спам. Эти клиенты и так часто приходят.

Можно уделять меньше внимания знакомству с брендом или продуктом, особенно если это товар широкого применения.

Можно создать реферальную программу, которая привлекает клиентов, которые приводят новых клиентов, например, предлагая скидку друзьям и родственникам вашего клиента.

Проводить опросы - здесь сразу две возможности:

- а) демонстрация того, что вам небезразлично мнение клиентов
- б) по результатам можно будет понять, какие покупатели предпочитают совершать покупки, сможем уточнить свое представление об идеальной целевой аудитории и сформировать акции с наибольшей вероятностью конверсии.

Сделать улучшения: если опрос выявит области, которые создают проблемы для клиентов, такие как неудобства при покупках, потребности по ассортименту, недовольство работой сотрудников аптеки, можно устранить их и выполнить глобальные фокусы (увеличить частоту посещения аптеки и увеличить средний чек)