

## RFM-анализ

- Что такое и для чего нужен RFM-анализ
- Как делать RFM-анализ
- Бизнес-кейс
- Выводы

### ➤ Что такое и для чего нужен RFM-анализ

RFM-анализ позволяет разделить клиентов на сегменты, основываясь на данных об их покупках. Он учитывает три признака: **recency** (давность покупки), **frequency** (частота покупок) и **monetary** (сумма трат клиента).

- Recency — этот показатель демонстрирует, как давно клиенты приобретали наши товары или пользовались нашими услугами;
- Frequency — с какой периодичностью клиенты обращаются за нашими товарами или услугами или сколько всего было приобретено товаров или услуг;
- Monetary — сколько всего у нас тратят.

На основе анализа можно выстраивать дальнейшую работу с клиентом; например, в зависимости от сегмента, куда попадает клиент, давать или не давать дополнительную скидку в рассылке.

Вот несколько примеров того, как использовать сегменты:

- чтобы поддерживать лояльность ключевых клиентов, нужно регулярно сообщать им про новинки и актуальные акции;
- клиентов, которые купили очень давно один раз, реактивировать через спецпредложение или промокод;
- тех, кто покупает редко, но на большую сумму, мотивировать покупать чаще большой суммой бонусов;
- «старых» клиентов реактивировать с помощью таргетированной рекламы.

Так, с помощью RFM-анализа можно сократить объем выдаваемых скидок и бонусов. Тем, кто постоянно покупает, необязательно предлагать дополнительный промокод, а «старым» скидка поможет вернуться к покупкам.

Частота проведения RFM-анализа очень сильно зависит от масштабов бизнеса. Так, крупные кампании могут пересчитывать сегменты RFM практически каждый день.

### ➤ Как делать RFM-анализ

Для проведения RFM-анализа нужно собрать данные обо всех покупках. По каждому из трёх показателей клиентскую базу делят на три группы.

Давность покупки	Частоты покупок	Сумма покупок
1 – недавние	1 – часто	1 – большая
2 – средняя давность	2 – средняя частота	2 – средняя
3 – давние	3 – редко	3 – маленькая

Единых правил для определения границ нет. Все зависит от самого бизнеса и его специфики.

Например, приобретение автомобиля год назад считается недавней покупкой, а приобретение продуктов в магазине два месяца назад – давно. Покупка одежды каждый месяц – это часто, покупка продуктов раз в месяц – редко.

Чтобы определить критерии распределения клиента в ту или иную группу обычно используют процентиля. Например, 25 и 75 или 33 и 66. Можно выбирать любые из, главное добиться того, чтобы клиенты были распределены равномерно по группам.

Дальше происходит сегментация.

### ➤ Бизнес-кейс

RFM-анализ был проведен на базе аптечной сети.

Для начала для каждого клиента (идентификатором был номер карты) было рассчитано:

- сколько дней прошло после последней покупки;
- сколько всего покупок было сделано;
- на какую сумму было совершено покупок.

По условию, если в момент покупки касса была в оффлайн-режиме, то вместо номера карты записывается зашифрованная последовательность символов. Поэтому карты были отфильтрованы, чтобы получить только идентифицированные транзакции.

Код представлен ниже на скриншоте:

```
with recency as (
    select card,
           min(day_diff) as day_after_last_purchase
    from (
        select card,
               date_part('day', last_purchase-datetime) as day_diff
        from (
            select card,
                   datetime,
                   doc_id,
                   max(datetime) over() as last_purchase -- была взята последняя дата покупки, которая есть в базе
            from bonuscheques b
            where length(card) = 13
        ) t1
        ) t2
    group by card
), -- расчет сколько дней прошло после последней покупки клиента
frequency as (
    select card,
           count(doc_id) as cnt_purchase
    from bonuscheques b
    where length(card) = 13
    group by card
), -- расчет количества покупок клиента
monetary as (
    select card,
           sum(summ_with_disc) as total_amount
    from bonuscheques b
    where length(card) = 13
    group by 1
), -- расчет на какую сумму было приобретено товаров каждым клиентом
```

После расчетов я посмотрел описательную статистику получившихся данных, их можно увидеть на скриншоте ниже:

- day\_after\_last\_purchase – количество прошедших дней после последней покупки;
- cnt\_purchase – количество покупок;
- total\_amount – на какую сумму было совершено покупок.

	day_after_last_purchase	cnt_purchase	total_amount
count	5926.000000	5926.000000	5926.000000
mean	113.374620	3.556362	3239.426257
std	95.150418	6.253204	6697.171918
min	0.000000	1.000000	21.000000
25%	30.000000	1.000000	729.000000
50%	87.000000	2.000000	1475.500000
75%	183.000000	4.000000	3315.750000
max	332.000000	217.000000	163943.000000

Далее, для сегментации, я рассчитал 25-й и 75 процентиля и получил такие данные, как показано на скриншоте ниже:

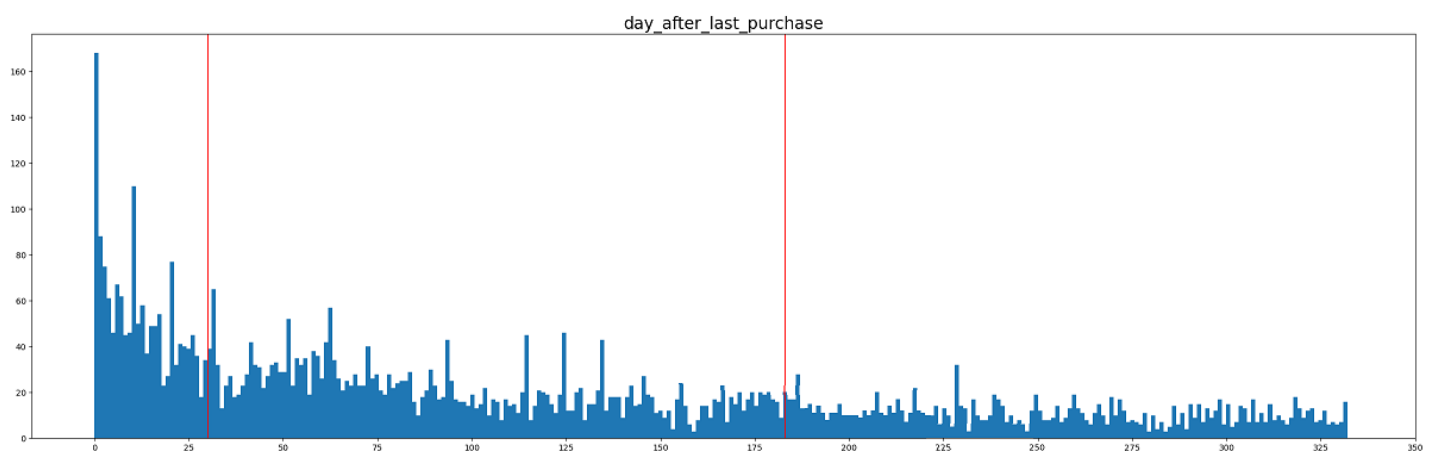
Для recency:  
 25-й процентиль: 30.0;  
 75-й процентиль: 183.0

Для frequency:  
 25-й процентиль: 1.0;  
 75-й процентиль: 4.0

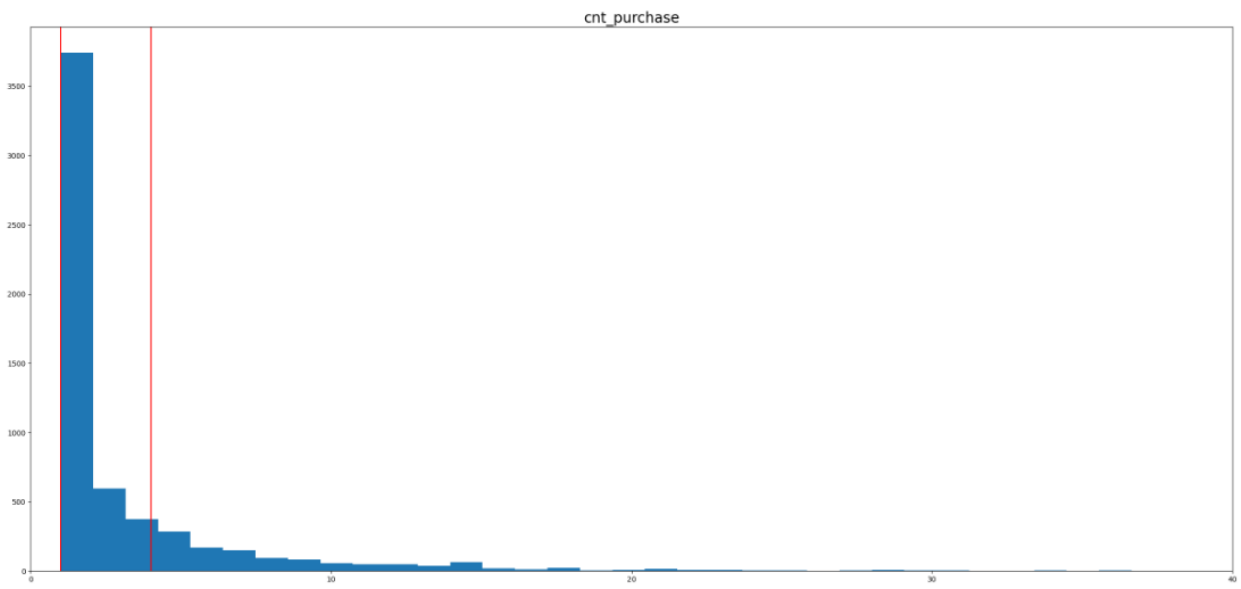
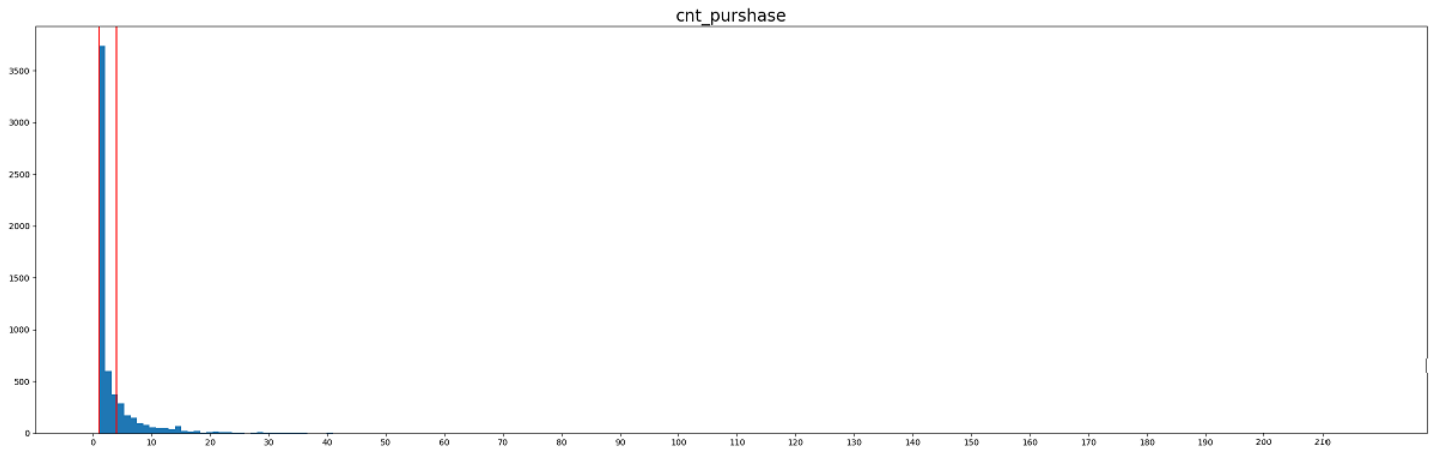
Для monetary:  
 25-й процентиль: 729.0;  
 75-й процентиль: 3315.75

Чтобы было нагляднее построил гистограммы и отметил процентиля:

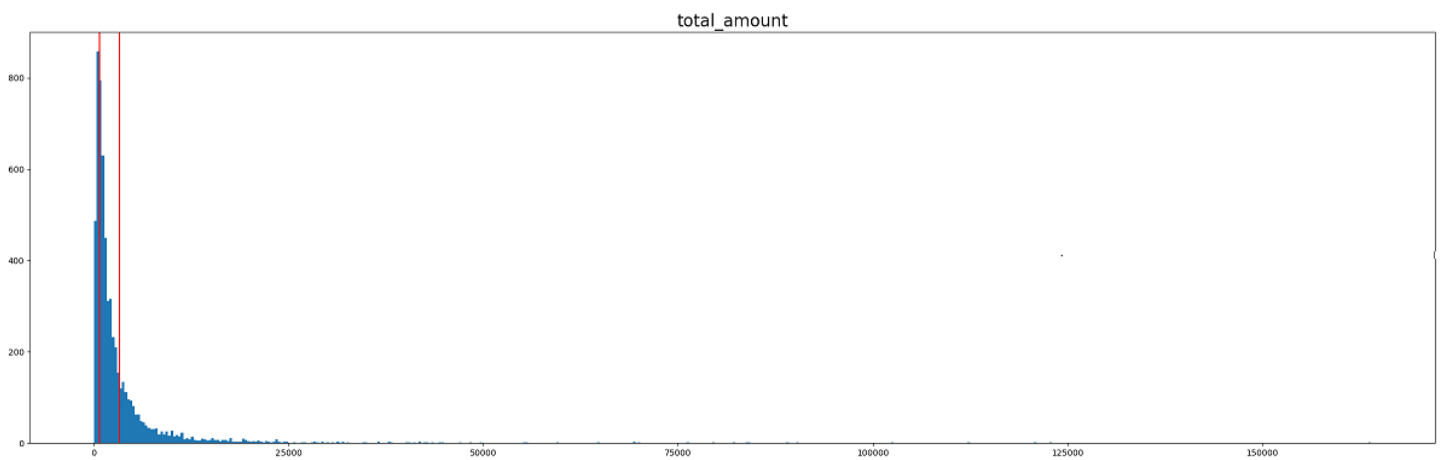
- Гистограмма по recency

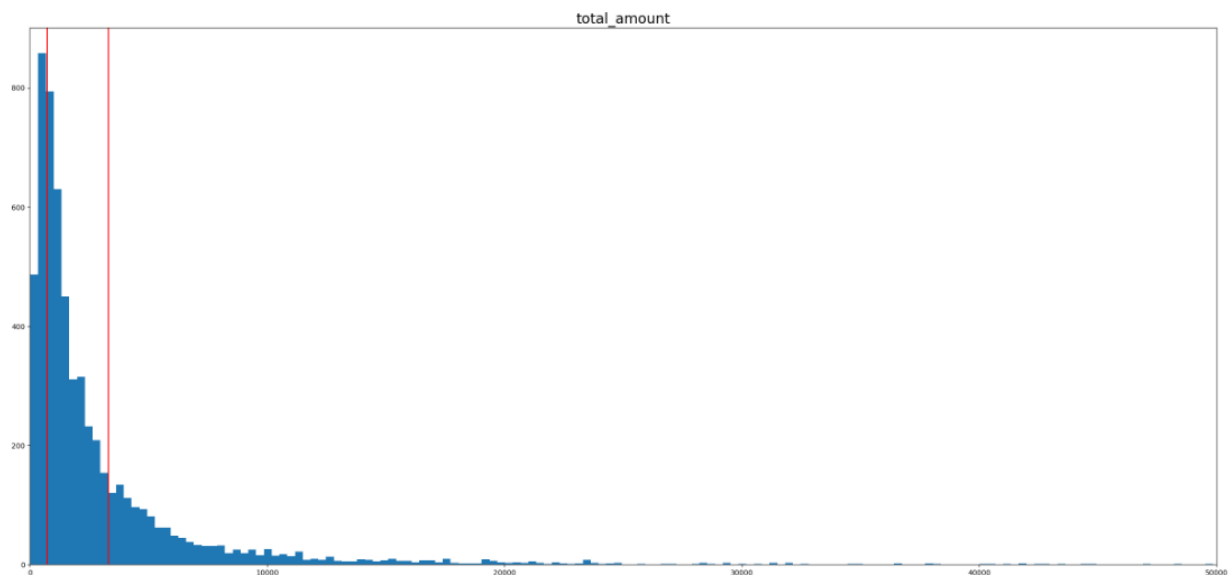


- Гистограмма по frequency



- Гистограмма по monetary





По гистограммам можно понять, что у нас присутствуют экстремальные значения в количестве покупок и сумме, на которую были совершены покупки, но для данного анализа убирать их будет не совсем корректно, т.к это просто постоянные клиенты, которые совершали покупки на всем рассматриваемом промежутке времени.

После расчета процентилей я сегментировал клиентов по группам и посчитал количество клиентов в каждой группе и их долю.

Группы по давности покупки:

	recency_group	cnt_clients_in_group	share, %
0	recency_1	1516	25.58
1	recency_2	2943	49.66
2	recency_3	1467	24.76

Группы по количеству покупок:

	frequency_group	cnt_clients_in_group	share, %
0	frequency_1	2686	45.33
1	frequency_2	2026	34.19
2	frequency_3	1214	20.49

Группы по сумме покупок:

	monetary_group	cnt_clients_in_group	share, %
0	monetary_1	1485	25.06
1	monetary_2	2959	49.93
2	monetary_3	1482	25.01

Группы получились не совсем равными, поэтому я решил посчитать еще 33-й и 66-й процентиля, результаты видно на скриншоте ниже:

```
Для recency:
33-й процентиль: 48.0;
66-й процентиль: 143.0
```

---

```
Для frequency:
33-й процентиль: 1.0;
66-й процентиль: 3.0
```

---

```
Для monetary:
33-й процентиль: 922.0;
66-й процентиль: 2403.5
```

И также сегментировал клиентов по группам, исходя из новых пороговых значений, которые являются процентилями.

Группы по давности покупки:

	recency_group	cnt_clients_in_group	share, %
0	recency_1	1983	33.46
1	recency_2	1940	32.74
2	recency_3	2003	33.80

Группы по количеству покупок:

	frequency_group	cnt_clients_in_group	share, %
0	frequency_1	2686	45.33
1	frequency_2	1652	27.88
2	frequency_3	1588	26.80

Группы по сумме покупок:

	monetary_group	cnt_clients_in_group	share, %
0	monetary_1	1957	33.02
1	monetary_2	1954	32.97
2	monetary_3	2015	34.00

С 33-им и 66-ым процентилями клиенты стали более равномерно распределены по группам, поэтому оставим данные пороги.

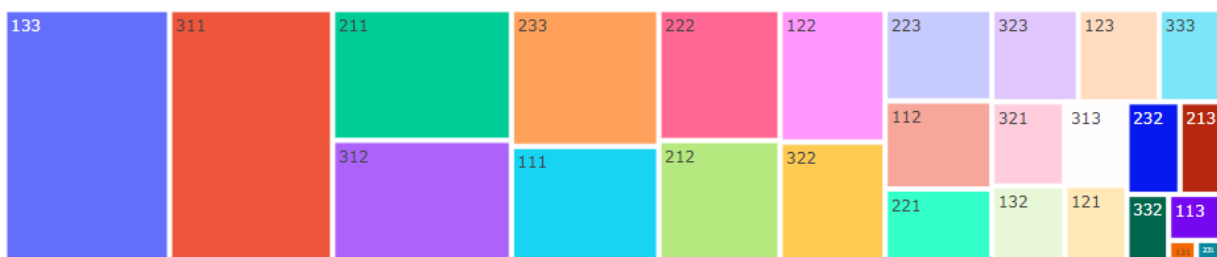
## Расчет процентилей в SQL:

```
all_card as (
    select distinct card
    from bonuscheques
    where length(card) = 13
), --все клиенты
pivot_table as (
    select ac.card as client,
           day_after_last_purchase,
           cnt_purchase,
           total_amount
    from all_card ac
    left join recency r
    on r.card = ac.card
    left join frequency f
    on f.card = ac.card
    left join monetary m
    on m.card = ac.card
),
percentiles as (
    select percentile_cont(0.33) WITHIN GROUP (ORDER BY day_after_last_purchase) as recency_33,
           percentile_cont(0.66) WITHIN GROUP (ORDER BY day_after_last_purchase) as recency_66,
           percentile_cont(0.33) WITHIN GROUP (ORDER BY cnt_purchase) as frequency_33,
           percentile_cont(0.66) WITHIN GROUP (ORDER BY cnt_purchase) as frequency_66,
           percentile_cont(0.33) WITHIN GROUP (ORDER BY total_amount) as monetary_33,
           percentile_cont(0.66) WITHIN GROUP (ORDER BY total_amount) as monetary_66
    from pivot_table
) --расчет процентилей для сегментации клиентов
```

## Сегментация клиентов в SQL:

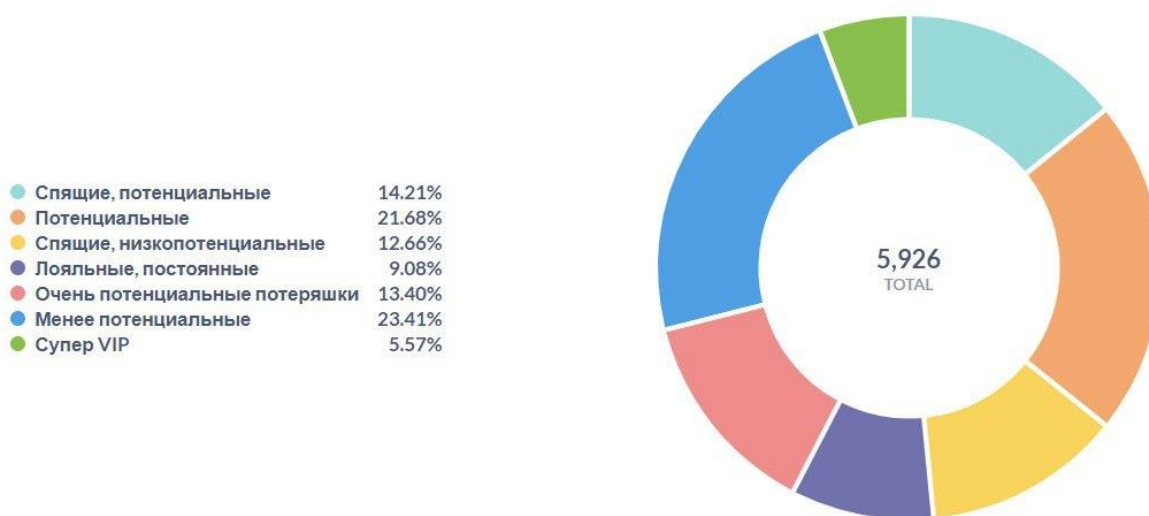
```
select client,
       RFM_segment,
       (case when RFM_segment in ('112', '121', '122') then 'Лояльные, постоянные'
       when RFM_segment in ('211', '212', '221', '222', '231', '232')
       then 'Потенциальные'
       when RFM_segment in ('113', '123', '131', '132', '133', '213',
       '223') then 'Менее потенциальные'
       when RFM_segment in ('312', '321', '322', '331', '332') then
       'Спящие, потенциальные'
       when RFM_segment in ('111') then 'Супер VIP'
       when RFM_segment in ('311') then 'Очень потенциальные потеряшки'
       else 'Спящие, низкопотенциальные' end) as "Название
       группы"
from (
    select client,
           concat(recency, frequency, monetary) as RFM_segment
    from (
        select client,
               case
                   when day_after_last_purchase <= recency_33
                   then 1
                   when day_after_last_purchase <= recency_66
                   then 2
                   else 3
               end as recency,
               case
                   when cnt_purchase <= frequency_33
                   then 1
                   when cnt_purchase <= frequency_66
                   then 2
                   else 3
               end as frequency,
               case
                   when total_amount <= monetary_33
                   then 1
                   when total_amount <= monetary_66
                   then 2
                   else 3
               end as monetary
        from pivot_table, percentiles
    ) main
) main_2
```

Исходя из каждой группы RFM получим итоговый сегмент каждого клиента и построим карту дерева:



В результате мы получили сегментированные группы клиентов, с которыми можно вести определенную коммуникацию.

### ➤ Выводы



### Спящие потенциальные

В эту группу я отнес клиентов сегментов 312, 321, 322, 331, 332.

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

Этим клиентам нужно про нас напомнить:

- Смс-кой или пушом „Начисленные бонусы с предыдущих покупок скоро сгорят/сгорят до X дня”;
- рассказать про новинки или скидки.

### Потенциальные

В эту группу я отнес клиентов сегментов 211, 212, 221, 222, 231, 232.

Платежеспособные клиенты, которые покупают нерегулярно. Основная наша цель сделать так, чтобы они покупали чаще.

Придумать акции на количество покупок за, например, месяц, чтобы была больше скидка на следующий период или повышенный кешбэк за более частые покупки.

Также напоминать про новинки, акции.



### **Лояльные постоянные**

В эту группу я отнес клиентов сегментов 112, 121, 122.

Эта группа клиентов – группа лидеров. Они уже приносят основной доход и покупают часто. С этой группой можно работать над повышением среднего чека.

Можно приглашать подписаться на соцсети, проводить опросы.

Не стоит делать частую рассылку про скидки, новинки и так далее; они и так покупают регулярно.

Для этой группы может очень хорошо зайти реферальная программа.

### **Очень потенциальные потеряшки**

В эту группу я отнес клиентов сегмента 311.

Клиенты этой группы покупали на очень хорошие суммы и регулярно, но давно.

С этой группой стоит поработать, так как они могут стать нашими супер VIP.

Стоит узнать почему они давно ничего не приобретали, постараться их реактивировать. Например, сделать разовую очень хорошую скидку или акцию, чтобы они вспомнили/вернулись к нам.

### **Менее потенциальные**

В эту группу я отнес клиентов сегментов 113, 123, 131, 132, 133, 213.

Данные клиенты покупают на меньшие деньги, чем потенциальные.

Этим клиентам может хорошо подойти такая акция, как „1+1=3”, начисления больше бонусов за большие покупки, скидки.

### **Супер VIP**

В эту группу я отнес клиентов сегмента 111.

Это абсолютные лидеры.

Тут уже стоит подумать над персонализированными акциями, скидками, бонусными картами. Тем, чего у других точно нет.

Это наши главные „фанаты”, а мы, конечно же, их.

### **Спящие низкопотенциальные**

В эту группу попали все остальные.

Давние, разовые или покупки на совсем небольшие суммы.

Не очень перспективная группа клиентов. Бюджета может уйти много, а профита быть совсем мало.

Также будет сложно определить точную причину почему данный клиент в этой группе.

Можно попробовать прислать смс-ку, чтобы напомнить о себе, рассказать, что можно, например, сделать заказ онлайн или открылась новая точка по новому адресу. Или предложить небольшую разовую акцию.