**RFM-анализ клиентской базы аптечной сети**

**Цель исследования:**  
Скоро сезон спадет и выручки в аптеке начнут падать. В компании есть механизм СМС-рассылок, но они не хотят тратить бюджеты впустую - они хотят провести персонализированную рассылку.

**Ключевые задачи:**

1. Провести классификацию пользователей
2. Сформировать рекомендации - какой группе пользователей что предлагать, чтобы увеличить количество покупок и средний чек

Для решения поставленной задачи отлично подойдет RFM-анализ.

Это метод анализа, позволяющий сегментировать клиентов по давности покупок (Recency), частоте (Frequency) и сумме (Monetary) покупок и выявлять тех клиентов, которые часто возвращаются и тратят много, и тех, кто заходил случайно и когда-то давно (их не стоит включать в рассылку).

По каждому показателю мы разделим пользователей на 3 группы:

1. Recency:  
   1 – был недавно  
   2 – был какое-то время назад  
   3 – был очень давно
2. Frequency:

1 – покупал очень часто  
2 – покупал регулярно

3 – покупал несколько раз

1. Monetary:

1 – большие суммы

2 – средние суммы

3 – маленькие суммы

Например, клиент из группы 123 – был недавно, покупает у нас регулярно, но при этом тратит немного.

В итоге максимум может получиться 27 групп.

**Структура базы данных**

Для исследования использовалась таблица **Bonuscheques**, которая содержит информацию о транзакциях по бонусным картам **card** (т.е. это клиенты, которые 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 «случайные» покупатели, сделавшие всего одну покупку.

**datetime** - дата и время совершения транзакции

**summ** - сумма чека

**Выбор метрик для анализа**

Основная трудность при проведении RFM-анализа - подобрать оптимальные пороги. Например, каких клиентов относить в группу 1 по признаку Recency, а кого в группу 2 или 3. Поэтому для этого исследования были рассчитаны среднее и медиана по каждому показателю.

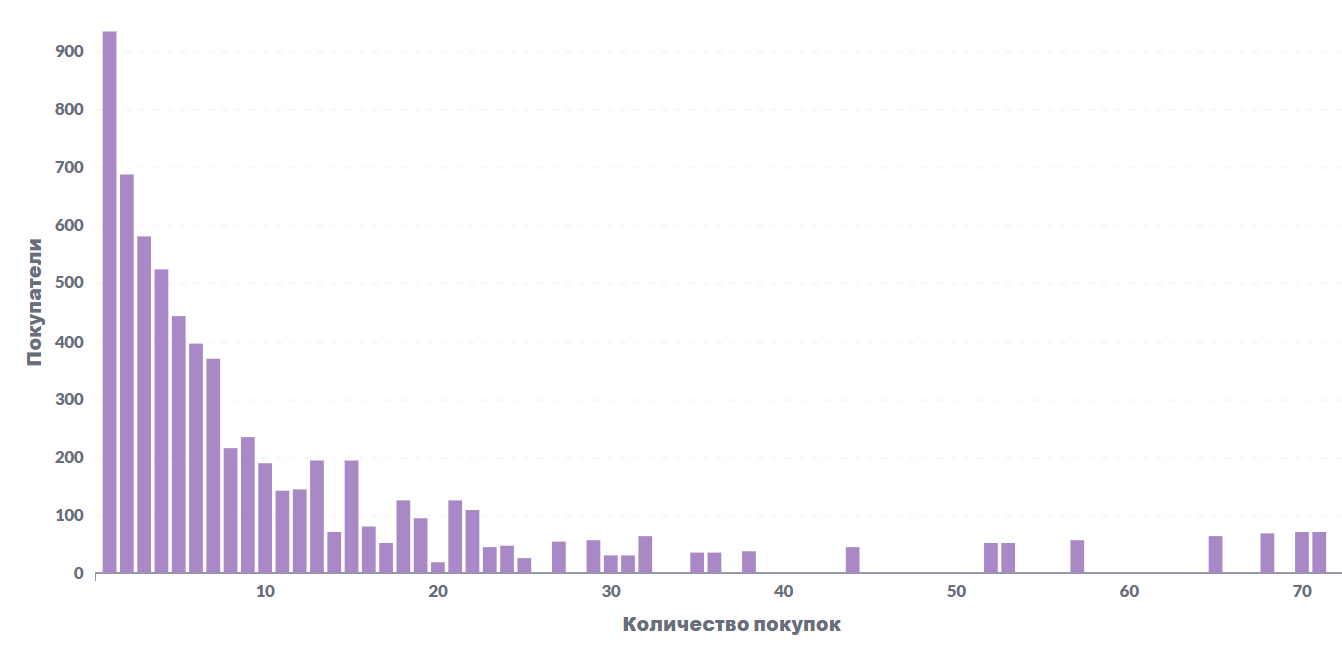
Среднее – один из самых популярных показателей, но он очень чувствителен к выбросам, поэтому чтобы увидеть более полную картину, определяем срединное значение(медиану) в группах.

Среднее и медиана по давности посещения не сильно разнятся (148,6 и 140 дней соответственно). Это может быть связано с тем, что у аптек много новых клиентов, которые были совсем недавно и так же, большое количество клиентов, которые очень давно не приходили. Поэтому не стоит брать слишком большой промежуток для анализа, чтобы не включать в рассылку “потерянных” клиентов.

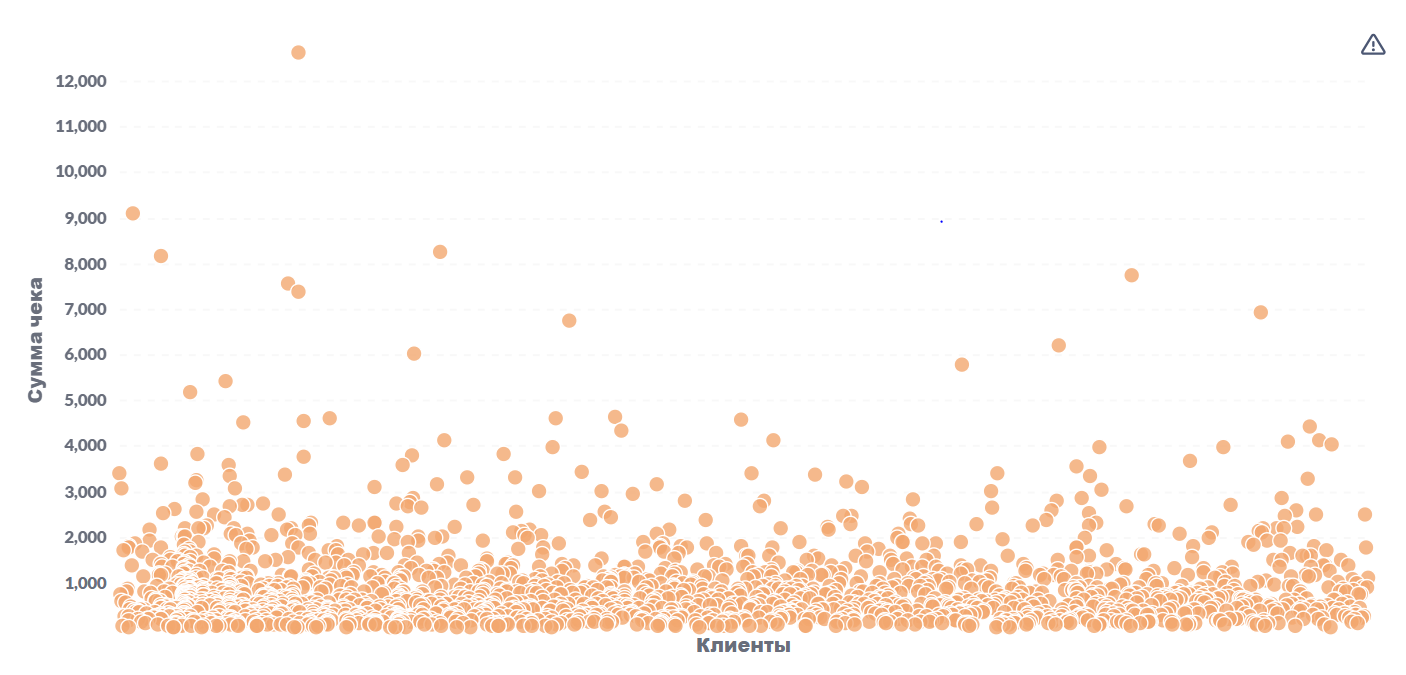
На графике (рис. 1.) видно, что основная часть покупателей приходит не больше 7 раз. Рассчитанное среднее количество посещений аптек 5,7, а медиана 3, что говорит о том, что у аптеки много разовых клиентов и нужно развивать систему лояльности, чтобы возвращать клиентов и сделать их постоянными.

Средняя сумма покупки составляет 5018, а медиана 2708, что почти в 2 раза меньше среднего, это говорит о том, что некоторые, единичные клиенты покупают на большие суммы, это можно увидеть на графике (рис.2).

**Рис.1. График количества покупок**



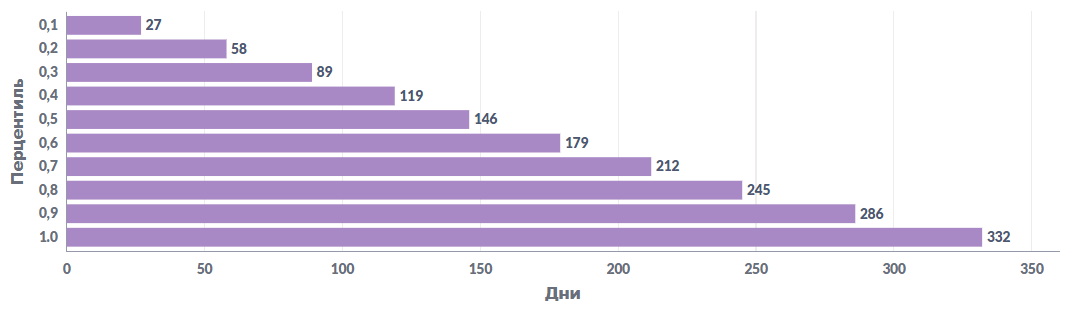
**Рис. 2. График распределения суммы покупок.**



**Деление на группы**

**Recency**

**Рис. 3. Распределение по количеству дней с момента последней покупки**



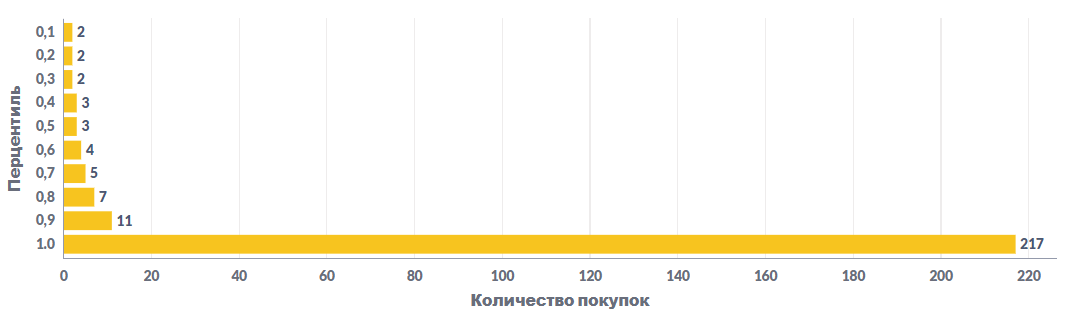
Чтобы разделить группы на 3 одинаковых части, в итоговом расчете берем 33й и 66й перцентиль.

Для признака Recency это 92 и 193 день, то есть, клиенты, которые приходили в аптеку менее чем 92 дня назад можем считать недавними и относим в группу 1. Во 2ю группу попадают среднее и медиана. И в группу 3 попадают те, кто был более чем 193 дня назад.

**Frequency**

Для этого показателя 33й и 66й перцентиль равны 5 и 3 соответственно, то есть, клиенты, которые приходили в сеть аптек более 5 раз мы относим в категорию 1, в нее попадает среднее число покупок. Те, кто покупал 4-5 раз попадают в группу 2. В группе 3 большинство клиентов может быть «случайными», которые зашли всего 1-2 раза, и также, там могут быть новые клиенты (кто заходил недавно), которым стоит уделить внимание и попытаться сделать их постоянными.

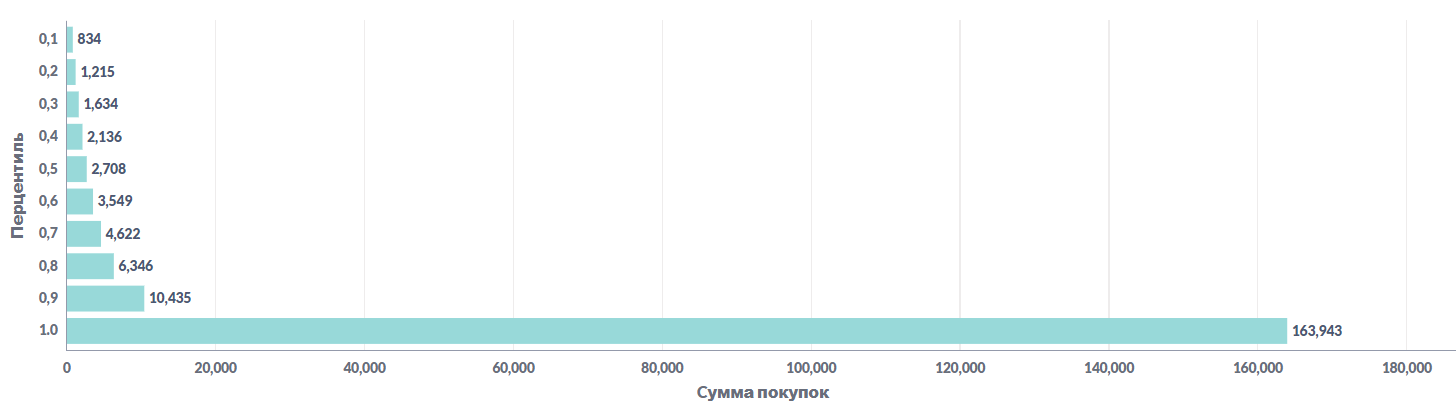
**Рис. 4. Распределение по количеству покупок на одного клиента**



**Monetary**   
Для этого показателя берем границы групп 4267 и 1812 рублей. В первую группу относим тех, кто покупал на сумму больше 4267, вместе с категорией 2 они приносят 90% выручки (накопительный итог). Оставшиеся 10% это категория 3, кто покупал на небольшие и скорее всего не часто.

На графике (рис.5) наглядно видно, что некоторые из клиентов покупали на большие суммы, что сильно исказило представление о сумме средней покупки.

**Рис. 5. Распределение по сумме покупок на одного клиента**



**Цель разделения на группы:** улучшить понимание покупательского поведения, разработать эффективные стратегии удержания клиентов и перевода их в более высокую группу (спящие потенциальные -> потенциальные -> постоянные), а также повысить их лояльность. Для этого нужно разработать персонализированные предложения, чтобы увеличить частоту покупок, средний чек и доход компании.

**Какие гипотезы можно выдвинуть?**

1. **Клиенты с недавними покупками более восприимчивы к маркетинговым кампаниям.** Те, кто недавно совершил покупку, с большей вероятностью откликнутся на предложения и акции, так как они находятся в активной фазе взаимодействия с сетью аптек.
2. **Клиенты, совершающие более крупные покупки, более ценны для бизнеса.** Клиенты, которые тратят больше на покупки, являются основными источниками дохода и должны рассматриваться как приоритетные для программ лояльности и персонализированных предложений.
3. **Разовые покупатели могут нуждаться в дополнительной мотивации для повторных покупок.** Клиенты, которые покупали только один раз, могут не иметь устойчивого интереса к сети аптек и требуют специальных предложений или стимулирования для их вовлечения и удержания.
4. **Давние клиенты с регулярными покупками имеют высокий потенциал для кросс-продаж.** Клиенты, которые покупают часто и на большие суммы, могут быть более склонны к кросс-продажам других продуктов, включая витамины, добавки или средства для ухода.

**SQL-запрос**

Этот код был написан для СУБД PostgreSQL с небольшой модификацией для создания фильтрации при построении дашборда в Metabase.

**with** *orders\_date* **as** (

**select**

card,

datetime::**date** **as** *order\_date*,

to\_date(**'2022-06-09'**, **'YYYY-MM-DD'**) - datetime::**date** **as** *diff*, -- '2022-06-09' - дата последней записи в БД

**row\_number**() **over** (**partition** **by** card) **as** *rn*

**from** bonuscheques

**where** card **like** **'200%'** -- берe данные карт, когда касса была в онлайн режиме

**and** {{datetime}} --фильтр даты для дашборда в Metabase

**order** **by** card, datetime **desc**

),

*recency* **as** (

**select** card, order\_date, diff, rn

**from** orders\_date *od*

**where** rn = 1

),

*rfm* **as** (

**select** *b*.card,

**count**(\*) **as** *frequency*,

**sum**(summ) **as** *monetary*,

**max**(*r*.diff) **as** *recency*

**from** bonuscheques *b*

**join** recency *r*

**on** *r*.card = *b*.card

**where** *b*.card **like** **'200%'**

**group** **by** *b*.card

**having** **count**(\*) > 1 -- отсеиваем "случайных" покупателей, чтобы сконцентрироваться на более потенциальных

),

*perc* **as** (

-- находим 33й и 66й перцентиль, чтобы разделить группы на равные части

**select**

**percentile\_cont**(0.33) **WITHIN** **GROUP** (**ORDER** **BY** recency) **as** *per\_recency1*,

**percentile\_cont**(0.66) **WITHIN** **GROUP** (**ORDER** **BY** recency) **as** *per\_recency2*,

**percentile\_cont**(0.33) **WITHIN** **GROUP** (**ORDER** **BY** frequency **desc**) **as** *per\_frequency1*,

**percentile\_cont**(0.66) **WITHIN** **GROUP** (**ORDER** **BY** frequency **desc**) **as** *per\_frequency2*,

**percentile\_cont**(0.33) **WITHIN** **GROUP** (**ORDER** **BY** monetary **desc**) **as** *per\_monetary1*,

**percentile\_cont**(0.66) **WITHIN** **GROUP** (**ORDER** **BY** monetary **desc**) **as** *per\_monetary2*

**from** rfm

),

*rfm\_groups* **as** (

**select**

card,

recency,

**case**

**when** recency < per\_recency1 **then** **'1'** -- разделяем клиентов на группы по каждому признаку

**when** recency < per\_recency2 **then** **'2'**

**else** **'3'**

**end** **as** *"R"*,

frequency,

**case**

**when** frequency > per\_frequency1 **then** **'1'**

**when** frequency > per\_frequency2 **then** **'2'**

**else** **'3'**

**end** **as** *"F"*,

monetary,

**case**

**when** monetary > per\_monetary1 **then** **'1'**

**when** monetary > per\_monetary2 **then** **'2'**

**else** **'3'**

**end** **as** *"M"*

**from** rfm, perc *p*

),

*rfm\_common* **as** (

**select**

card **as** *client*,

recency,

frequency,

monetary,

concat("R","F","M") **as** *"RFM"*

**from** rfm\_groups

**order** **by** *"RFM"*

),

*rfm\_tab* **as**(

**select**

client,

"RFM",

-- формирую и называю группы

**case**

**when** "RFM" **in** (**'111'**) **then** **'VIP клиенты'**

**when** "RFM" **in** (**'112'**, **'121'**, **'122'**) **then** **'Лояльные, постоянные'**

**when** "RFM" **in** (**'113'**, **'213'**, **'123'**) **then** **'Постоянные с небольшим бюджетом'**

**when** "RFM" **in** (**'221'**, **'122'**, **'131'**, **'212'**, **'231'**) **then** **'Потенциальные'**

**when** "RFM" **in** (**'312'**, **'313'**, **'321'**, **'322'**) **then** **'Спящие потенциальные'**

**when** "RFM" **in** (**'133'**, **'132'**, **'222'**, **'223'**, **'232'**, **'233'**) **then** **'Низкопотенциальные'**

**when** "RFM" **in** (**'211'**, **'311'**) **then** **'Потерянные VIP'**

**else** **'Спящие низкопотенциальные'**

**end** **as** *"Название группы"*

**from** rfm\_common

)

**select** -- создаю итоговую таблицу

orders\_date.card **as** *"Клиент"*,

**max**(order\_date) **as** *"Дата"*,

**max**(recency) **as** *"Дни с послед. покупки"*,

**max**("R") **as** *"R"*,

**max**(frequency) **as** *"Количество покупок"*,

**max**("F") **as** *"F"*,

**max**(monetary) **as** *"Сумма покупок"*,

**max**("M") **as** *"M"*,

**max**("RFM") **as** *"RFM"*,

**max**("Название группы") **as** *"Название группы"*

**from** orders\_date

**join** rfm\_tab

**on** *rfm\_tab*.client = orders\_date.card

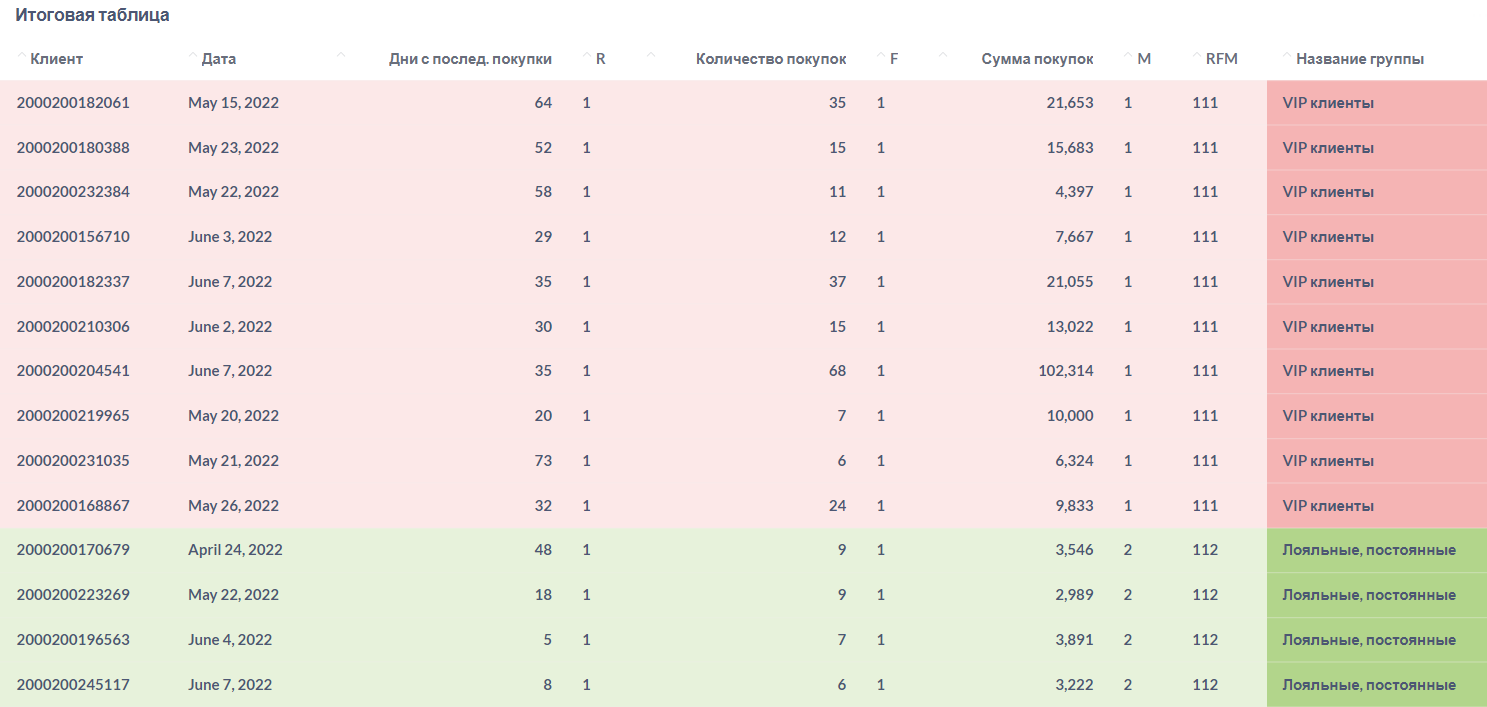
**join** rfm\_groups

**on** *rfm\_groups*.card = orders\_date.card

**group** **by** orders\_date.card

**order** **by** *"RFM"*

**Рис.6. Итоговая таблица RFM-анализа**

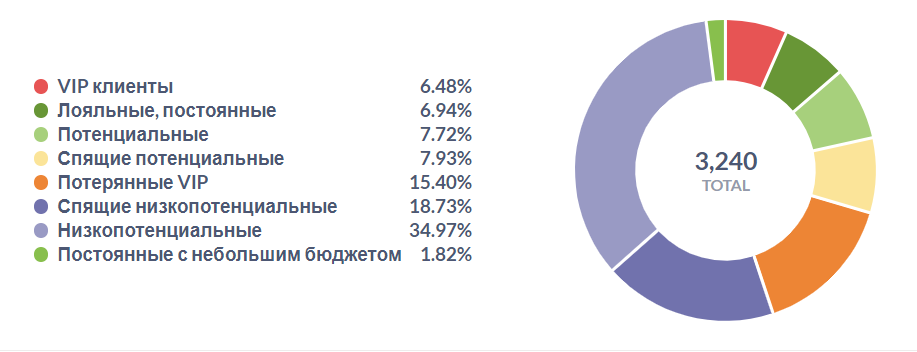


**Ссылка на дашборд:**<https://metabase.simulative.ru/dashboard/301-rfm-dashboard#refresh=600>

**Выводы и рекомендации по рассылке:**

В результате анализа в некоторых группах получилось совсем небольшое количество людей, поэтому их можно объединить в схожие смысловые группы. В итоге получилось 8 групп для смс-рассылки.

**Рис.7. Распределение клиентов по группам**



Основной фокус на том, чтобы увеличить количество покупок и средний чек. Это можно сделать, переведя клиентов на вышестоящий уровень воронки, т.е.:

Попытаться вернуть «потерянных» клиентов

Эмодзи ⬇️ Facebook.com

«Спящих» клиентов сделать потенциальными

Эмодзи ⬇️ Facebook.com

Потенциальных сделать лояльными

Эмодзи ⬇️ Facebook.com

Лояльных поднять до уровня VIP

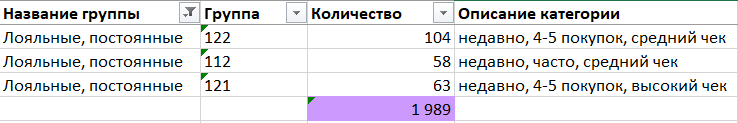
Эмодзи ⬇️ Facebook.com

VIP клиентам предложить такой сервис, чтобы они не только оставались постоянными клиентами, но и советовали друзьям и родственникам

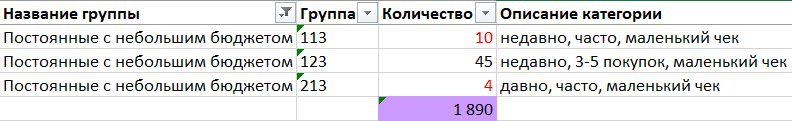
Группе «**VIP клиентов»** можно выпустить эксклюзивную карту, по которой будет начисляться кешбэк (если кешбэк уже предусмотрен для всех клиентов, то для VIP группы можно сделать повышенный кешбэк), чтобы они и дальше оставались клиентами данной сети.



**«Лояльные, постоянные»** - им можно разослать скидку на витамины, средства для ухода, так как у них уже есть доверие к аптеке, и они будут более склонны к покупке товаров из других категорий. Так же, доверие к аптеке обычно выше, чем к продавцам на маркетплейсам.



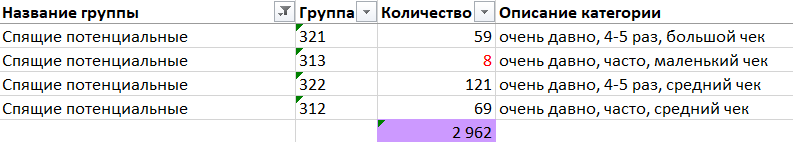
**«Постоянные с небольшим бюджетом»** - эта группа важна, так как даже небольшую, но регулярную прибыль приносит компании. Для них в рассылке будет актуальна акция в формате 2+1 или купи 3 и получи подарок.



**«Потенциальные»** - они заходят периодически в данную сеть аптек, но не очень часто, поэтому для них нужно сделать акцию, которая заставит их чаще покупать. Например, каждая 5я покупка со скидкой х% или собери 5 чеков/наклеек и получи подарок.

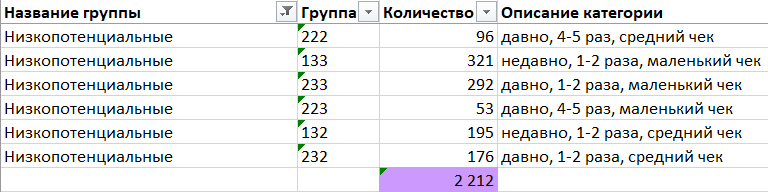


**«Спящие потенциальные»** - раньше они заходили часто, тратили в основном средние и большие суммы, поэтому им нужно напомнить о себе, чтобы сделать их снова сделать постоянными покупателями. Для этого им можно предложить скидку или двойные бонусы до определенного числа, напомнить, что у них остались бонусы на карте (если они есть).



**«Низкопотенциальные»** - в этой категории велик шанс, что выделенный бюджет на данную категорию может не окупиться, но в эту категорию могли попасть новые клиенты, которые недавно зашли всего 1 раз. Не плохо было бы их сделать постоянными клиентами, даже если они покупают на небольшие суммы, поэтому можно разослать напоминание и скидку:

«Здравствуйте, [Имя клиента]! Не забывайте заботиться о своем здоровье, а чтобы это было еще проще, дарим вам скидку х% на следующую покупку!»

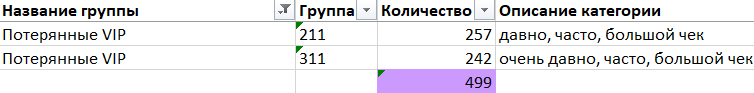


**«Потерянные VIP»** - эту категорию нужно обязательно попытаться вернуть и сделать постоянными клиентами, так как они раньше заходили часто и тратили большие суммы. Скидкой таких клиентов не всегда удается вернуть, ведь они и без скидок тратили достаточно много в сети аптек. Нужно показать, что компания заботится о них. Для этого можно просмотреть покупки, данной категории клиентов и сделать персонализированную рассылку, в которой напомнить о себе и показать, что компания их ждет.

Например, «Дорогой [Имя клиента]! Вы давно к нам не заходили, витамины, которые Вы у нас купили, должно быть уже закончились, мы подготовили новые для Вас, ждем вас в сети аптек «…»»

Если нет возможности сделать персонализированную рассылку, вот пример для рассылки группе «потерянные VIP»:

«Дорогой [Имя клиента], мы всегда рады видеть вас! Не забывайте заботиться о своем здоровье. Если вам что-то нужно, мы здесь, чтобы помочь!»



**«Спящие низкопотенциальные»** - на эту категорию не стоит возлагать больших надежд, потому что за долгое время отсутствия, они могли сменить место жительства, номер или уйти к конкурентам.

При рассылке сообщений этой категории большой риск потратить бюджет и не получить результата, потому что в этой категории те, кто заходил случайно, очень давно и покупали недорогие товары. Возможно, лучше стоит потратить бюджет на привлечение новых клиентов.

